Although the ribosomes of bacteria and eukaryotes are very similar in structure and function, eukaryotic ribosomes are slightly larger, as well as differing somewhat from bacterial ribosomes in their molecular composition. The differences are medically significant. Certain antibiotic drugs can inactivate bacterial ribosomes without affecting eukaryotic ribosomes. These drugs, including tetracycline and streptomycin, are used to combat bacterial infections.

The structure of a ribosome reflects its function of bringing mRNA together with tRNAs carrying amino acids. In addition to a binding site for mRNA, each ribosome has three binding sites for tRNA **(Figure 17.18)**. The **P site** (peptidyl-tRNA binding site) holds the tRNA carrying the growing polypeptide chain, while the **A site** (aminoacyl-tRNA binding site) holds the tRNA carrying the next amino acid to be added to the chain. Discharged tRNAs leave the ribosome from the **E site** (exit site). The ribosome holds the tRNA and mRNA in close proximity and positions the new amino acid so that it can be added to the carboxyl end of the growing polypeptide. It then catalyzes the formation of the peptide bond. As the polypeptide becomes longer, it passes through an *exit tunnel* in the ribosome's large subunit. When the polypeptide is complete, it is released through the exit tunnel.

The widely accepted model is that rRNAs, rather than ribosomal proteins, are primarily responsible for both the structure and the function of the ribosome. The proteins, which are largely on the exterior, support the shape changes of the rRNA molecules as they carry out catalysis during translation. Ribosomal RNA is the main constituent of the A and P sites and of the interface between the two subunits; it also acts as the catalyst of peptide bond formation. Thus, a ribosome could actually be considered one colossal ribozyme!

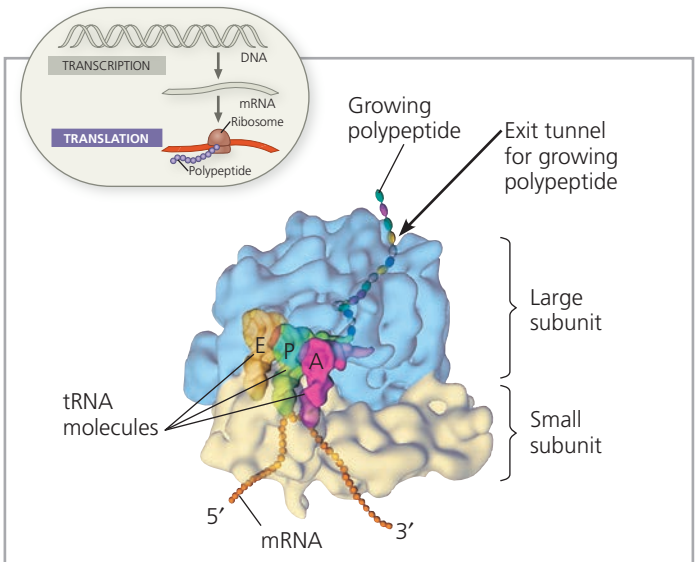
Building a Polypeptide

We can divide translation, the synthesis of a polypeptide, into three stages: initiation, elongation, and termination. All three require protein "factors" that aid in the translation process. Some steps of initiation and elongation also require energy, provided by the hydrolysis of guanosine triphosphate (GTP).

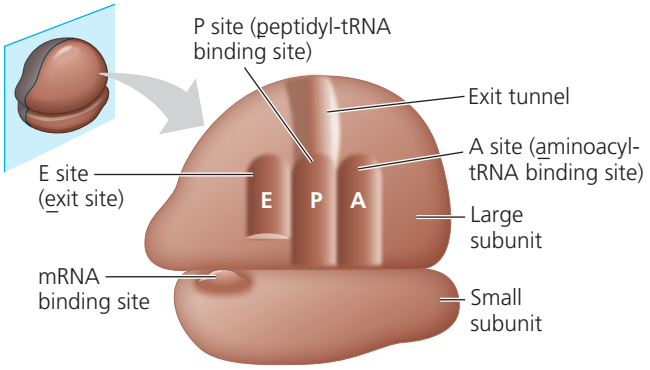
Ribosome Association and Initiation of Translation

In either bacteria or eukaryotes, the start codon (AUG) signals the start of translation; this is important because it establishes the codon reading frame for the mRNA. In the first step of translation, a small ribosomal subunit binds to both the mRNA and a specific initiator tRNA, which carries the amino acid methionine. In bacteria, the small subunit can bind the two in either order; it binds the mRNA at a specific RNA sequence, just upstream of the AUG start codon. In the **Scientific Skills Exercise**, you can work with DNA sequences encoding the ribosomal binding sites on the mRNAs of a group of *Eschericia coli*

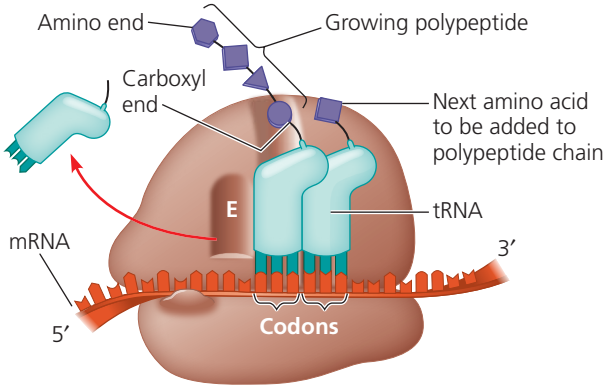
Figure 17.18 The anatomy of a functioning ribosome.



(a) Computer model of functioning ribosome. This is a model of a bacterial ribosome, showing its overall shape. The eukaryotic ribosome is roughly similar. A ribosomal subunit is a complex of ribosomal RNA molecules and proteins.



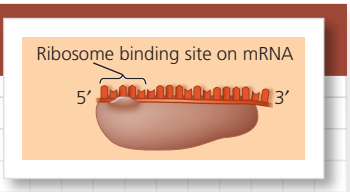
(b) Schematic model showing binding sites. A ribosome has an mRNA binding site and three tRNA binding sites, known as the A, P, and E sites. This schematic ribosome will appear in later diagrams.



(c) Schematic model with mRNA and tRNA. A tRNA fits into a binding site when its anticodon base-pairs with an mRNA codon. The P site holds the tRNA attached to the growing polypeptide. The A site holds the tRNA carrying the next amino acid to be added to the polypeptide chain. Discharged tRNAs leave from the E site. The polypeptide grows at its carboxyl end.

Animation: Ribosomes

SCIENTIFIC SKILLS EXERCISE



Interpreting a Sequence Logo

How Can a Sequence Logo Be Used to Identify Ribosome Binding Sites on Bacterial mRNAs? When initiating translation, ribosomes bind to an mRNA at a ribosome binding site upstream of the AUG start codon. Because mRNAs from different genes all bind to a ribosome, the genes encoding these mRNAs are likely to have a similar base sequence where the ribosomes bind. Therefore, candidate ribosome binding sites on mRNA can be identified by comparing DNA sequences (and thus the mRNA sequences) of multiple genes in a species, searching the region upstream of the start codon for shared (conserved) stretches of bases. In this exercise, you will analyze DNA sequences from multiple such genes, represented by a visual graphic called a sequence logo.

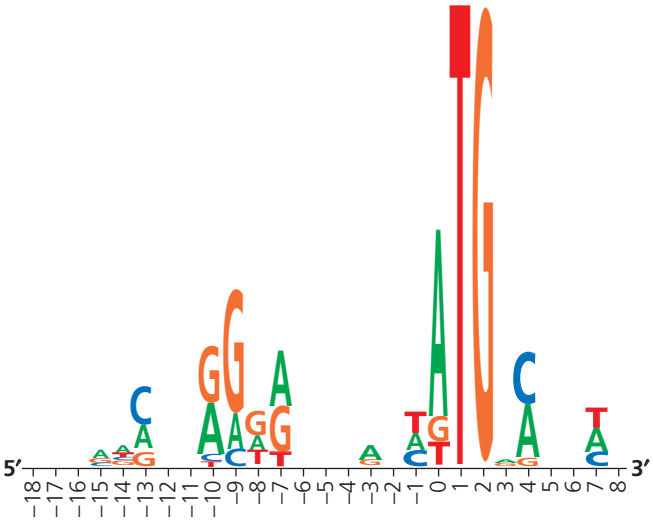
How the Experiment Was Done The DNA sequences of 149 genes from the *E. coli* genome were aligned using computer software. The aim was to identify similar base sequences—at the appropriate location in each gene—as potential ribosome binding sites. Rather than presenting the data as a series of 149 sequences aligned in a column (a sequence alignment), the researchers used a sequence logo.

Data from the Experiment To show how sequence logos are made, the potential ribosome binding regions from 10 *E. coli* genes are shown below in a sequence alignment, followed by the sequence logo derived from the aligned sequences. Note that the DNA shown is the nontemplate (coding) strand, which is how DNA sequences are typically presented.

	−18	−17	−16	−15	−14	−13	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8
thrA	G	G	T	A	A	C	G	A	G	G	T	A	A	C	A	A	C	C	A	T	G	C	G	A	G	T	G
lacA	C	A	T	A	A	C	G	G	A	G	T	G	A	T	C	G	C	A	T	T	G	A	A	C	A	T	G
lacY	C	G	C	G	T	A	A	G	G	A	A	A	T	C	C	A	T	T	A	T	G	T	A	C	T	A	T
lacZ	T	T	C	A	C	A	C	A	G	G	A	A	A	C	A	G	C	T	A	T	G	A	C	C	A	T	G
lacI	C	A	A	T	T	C	A	G	G	G	T	G	G	T	G	A	A	T	G	T	G	A	A	A	C	C	A
recA	G	G	C	A	T	G	A	C	A	G	G	A	G	T	A	A	A	A	A	T	G	G	C	T	A	T	C
galR	A	C	C	C	A	C	T	A	A	G	G	T	A	T	T	T	T	C	A	T	G	G	C	G	A	C	C
metJ	A	A	G	A	G	G	A	T	T	A	A	G	T	A	T	C	T	C	A	T	G	G	C	T	G	A	A
lexA	A	T	A	C	A	C	C	C	A	G	G	G	G	C	G	G	A	A	T	G	A	A	A	A	G	C	G
trpR	T	A	A	C	A	A	T	G	G	C	G	A	C	A	T	A	T	T	A	T	G	G	C	C	C	A	A

5′ 3′

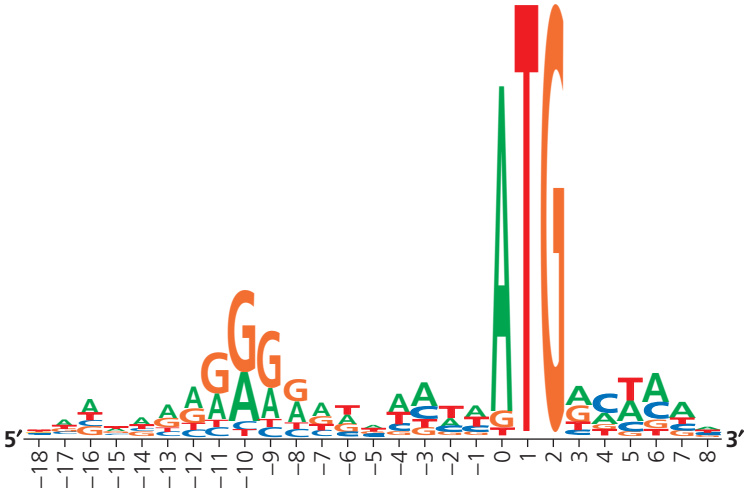
Sequence alignment



Sequence logo

INTERPRET THE DATA

1. In the sequence logo (bottom, left), the horizontal axis shows the primary sequence of the DNA by nucleotide position. Letters for each base are stacked on top of each other according to their relative frequency at that position among the aligned sequences, with the most common base as the largest letter at the top of the stack. The height of each letter represents the relative frequency of that base *at that position*. (a) In the sequence alignment, count the number of each base at position −9 and order them from most to least frequent. Compare this to the size and placement of each base at −9 in the logo. (b) Do the same for positions 0 and 1.
2. The height of a stack of letters in a logo indicates the predictive power of that stack (determined statistically). If the stack is tall, we can be more confident in predicting what base will be in that position if a new sequence is added to the logo. For example, at position 2 in the sequence alignment, all 10 sequences have a G; the probability of finding a G there in a new sequence is very high, as is the stack in the sequence logo. For short stacks, the bases all have about the same frequency, so it's hard to predict a base at those positions. (a) Looking at the sequence logo, which two positions have the most predictable bases? What bases do you predict would be at those positions in a newly sequenced gene? (b) Which 12 positions have the least predictable bases? How do you know? How does this reflect the relative frequencies of the bases shown at these positions in the sequence alignment? Use the two leftmost positions of the 12 as examples in your answer.
3. In the actual experiment, the researchers used 149 sequences to build their sequence logo, which is shown below. There is a stack at each position, even if short, because the sequence logo includes more data. (a) Which three positions in this sequence logo have the most predictable bases? Name the most frequent base at each. (b) Which four positions have the least predictable bases? How can you tell?



4. A consensus sequence identifies the base occurring most often at each position in the set of sequences. (a) Write out the consensus sequence of this (the nontemplate) strand. In any position where the base can't be determined, put a dash. (b) Which provides more information—the consensus sequence or the sequence logo? What is lost in the less informative method?
5. (a) Based on the logo, what five adjacent base positions in the 5′ UTR region are most likely to be involved in ribosome binding? Explain. (b) What is represented by the bases in positions 0–2?

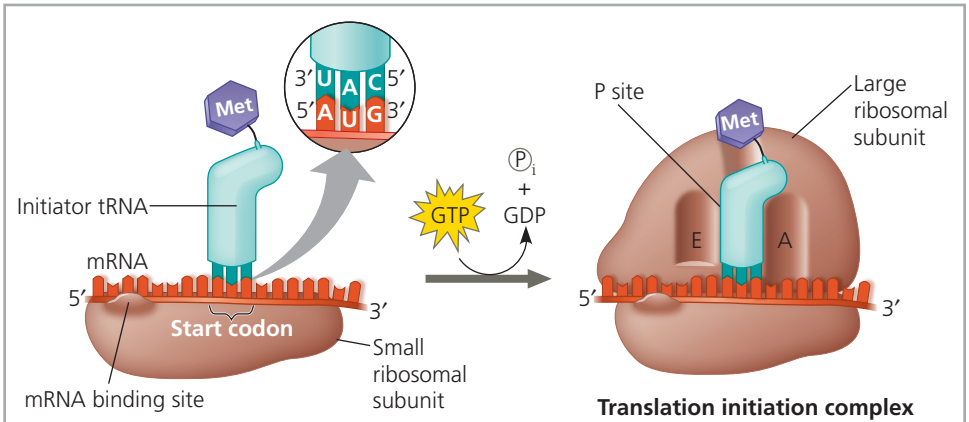
Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Further Reading T. D. Schneider and R. M. Stephens, Sequence logos: A new way to display consensus sequences, *Nucleic Acids Research* 18:6097–6100 (1990).

genes. In eukaryotes, the small subunit, with the initiator tRNA already bound, binds to the 5′ cap of the mRNA and then moves, or *scans*, downstream along the mRNA until it reaches the start codon; the initiator tRNA then hydrogen-bonds to the AUG start codon.

Thus, the first components to associate with each other during the initiation stage of translation are mRNA, a tRNA bearing the first amino acid of the polypeptide, and the small ribosomal subunit **(Figure 17.19)**. This is followed by the attachment of a large ribosomal subunit, completing the *translation initiation complex*. Proteins called *initiation factors* are required to bring all these components together. The cell also expends energy obtained by hydrolysis of a GTP molecule to form the initiation complex. At the completion of the initiation process, the initiator tRNA sits in the P site of the ribosome, and the vacant A site is ready for the next aminoacyl tRNA. Note that a polypeptide is always synthesized in one direction, from the initial methionine at the amino end, also called the N-terminus, toward the final amino acid at the carboxyl end, also called the C-terminus (see Figure 5.15).

Figure 17.19 The initiation of translation. Animation: Initiation of Translation



1 A small ribosomal subunit binds to a molecule of mRNA. In a bacterial cell, the mRNA binding site on this subunit recognizes a specific nucleotide sequence on the mRNA just upstream of the start codon. An initiator tRNA, with the anticodon UAC, base-pairs with the start codon, AUG. This tRNA carries the amino acid methionine (Met).

2 The arrival of a large ribosomal subunit completes the initiation complex. Proteins called initiation factors (not shown) are required to bring all the translation components together. Hydrolysis of GTP provides the energy for the assembly. The initiator tRNA is in the P site; the A site is available to the tRNA bearing the next amino acid.

Elongation of the Polypeptide Chain

In the elongation stage of translation, amino acids are added one by one to the previous amino acid at the C-terminus of the growing chain. Each addition involves several proteins called *elongation factors* and occurs in a three-step cycle described in **Figure 17.20**. Energy expenditure occurs in the first and third steps. Codon recognition requires hydrolysis of one molecule of GTP, which increases the accuracy and efficiency of this step. One more GTP is hydrolyzed to provide energy for the translocation step.

The mRNA is moved through the ribosome in one direction only, 5′ end first; this is equivalent to the ribosome moving 5′ → 3′ on the mRNA. The main point is that the ribosome and the mRNA move relative to each other, unidirectionally, codon by codon. The elongation cycle takes less than a tenth of a second in bacteria and is repeated as each amino acid is added until the polypeptide is complete. The empty tRNAs that are released from the E site return to the cytoplasm, where they will be reloaded with the appropriate amino acid (see Figure 17.17).

Termination of Translation

The final stage of translation is termination **(Figure 17.21)**. Elongation continues until a stop codon in the mRNA reaches the A site. The nucleotide base triplets UAG, UAA, and UGA (all written 5′ → 3′) do not code for amino acids but instead act as signals to stop translation. A *release factor*, a protein shaped like an aminoacyl tRNA, binds directly to the stop codon in the A site. The release factor causes the addition of a water molecule instead of an amino acid to the polypeptide chain. (Water molecules are abundant in the cytosol.) This reaction breaks (hydrolyzes) the bond between the completed polypeptide and the tRNA in the P site, releasing the polypeptide through the exit tunnel of the ribosome's large subunit. The remainder of the translation assembly then comes apart in a multistep process, aided by other protein factors. Breakdown of the translation assembly requires the hydrolysis of two more GTP molecules.

Completing and Targeting the Functional Protein

The process of translation is often not sufficient to make a functional protein. In this section, you will learn about modifications that polypeptide chains undergo after the translation process as well as some of the mechanisms used to target completed proteins to specific sites in the cell.

Protein Folding and Post-Translational Modifications

During its synthesis, a polypeptide chain begins to coil and fold spontaneously as a consequence of its amino acid sequence (primary structure), forming a protein with a specific shape: a three-dimensional molecule with secondary and tertiary structure (see Figure 5.18). Thus, a gene determines primary structure, which in turn determines shape.

Figure 17.20 The elongation cycle of translation. The hydrolysis of GTP plays an important role in the elongation process; elongation factors are not shown.

Animation: Elongation Cycle of Translation

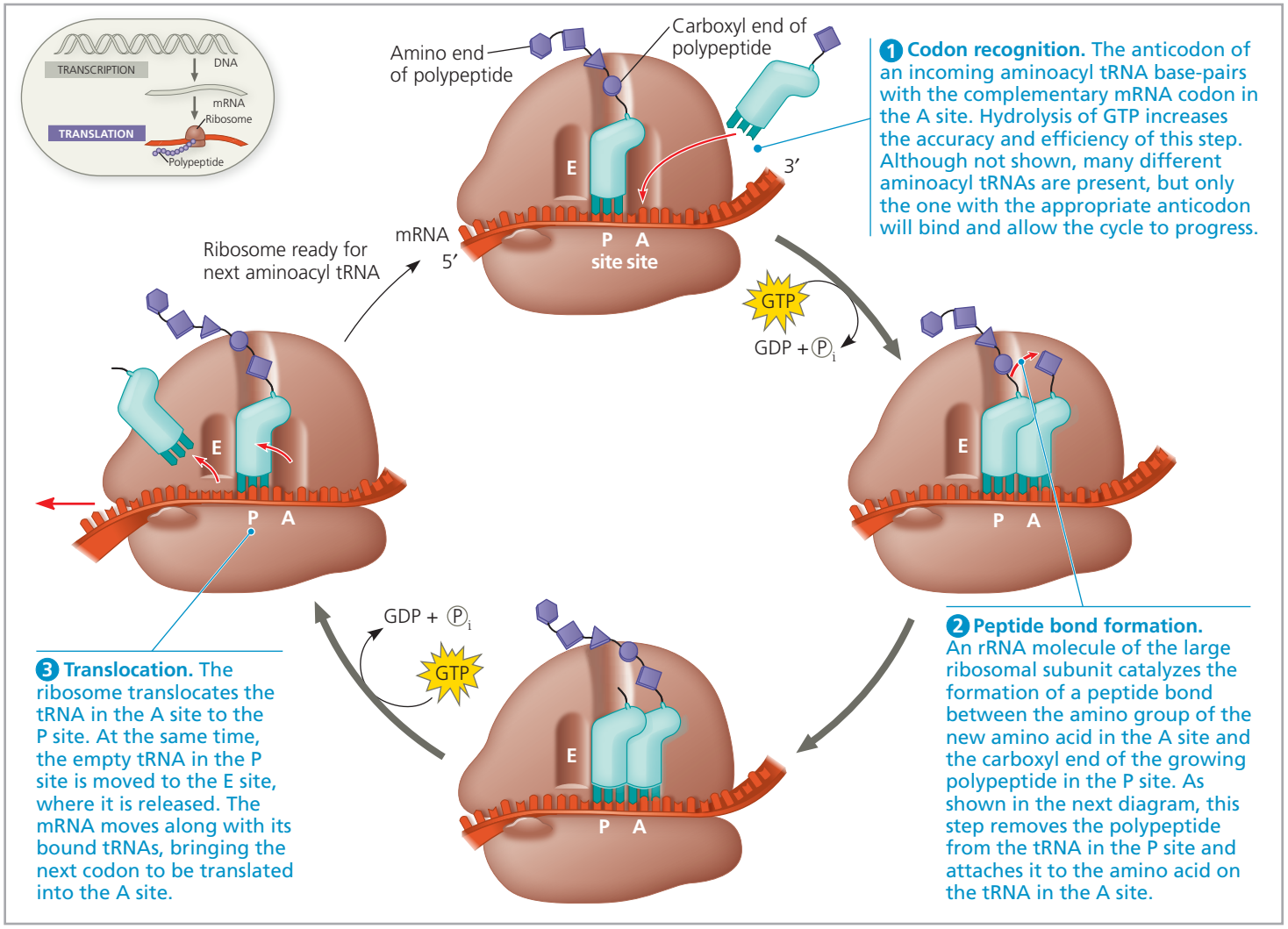
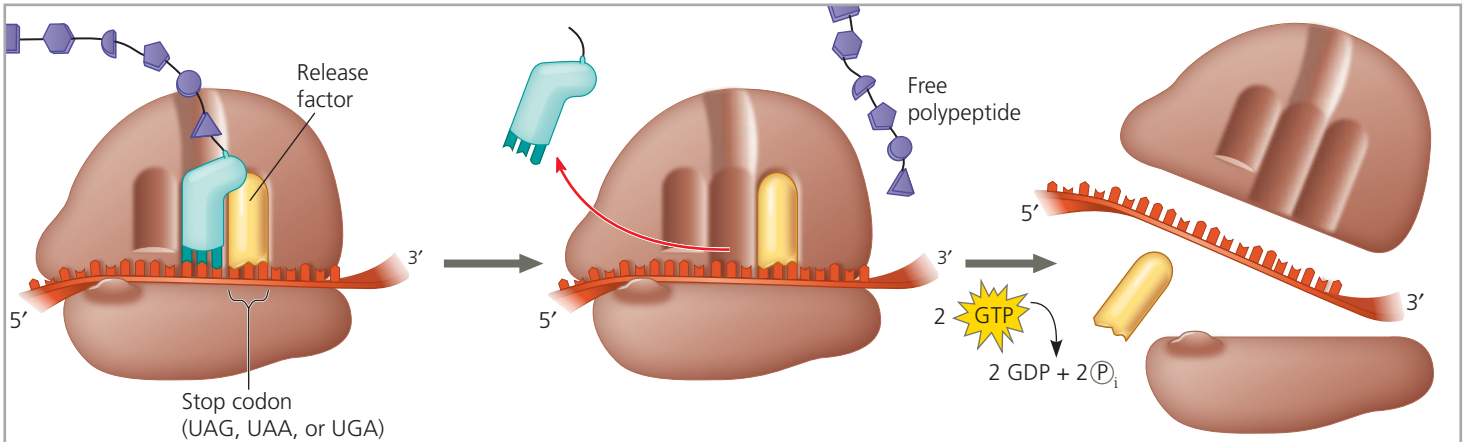


Figure 17.21 The termination of translation. Like elongation, termination requires GTP hydrolysis as well as additional protein factors, which are not shown here.

Animation: Termination of Translation



Additional steps—*post-translational modifications*—may be required before the protein can begin doing its particular job in the cell. Certain amino acids may be chemically modified by the attachment of sugars, lipids, phosphate groups, or other additions. Enzymes may remove one or more amino acids from the leading (amino) end of the polypeptide chain. In some cases, a polypeptide chain may be enzymatically cleaved into two or more pieces. In other cases, two or more polypeptides that are synthesized separately may come together, if the protein has quaternary structure; an example is hemoglobin (see Figure 5.18).

BioFlix® Animation: Protein Processing

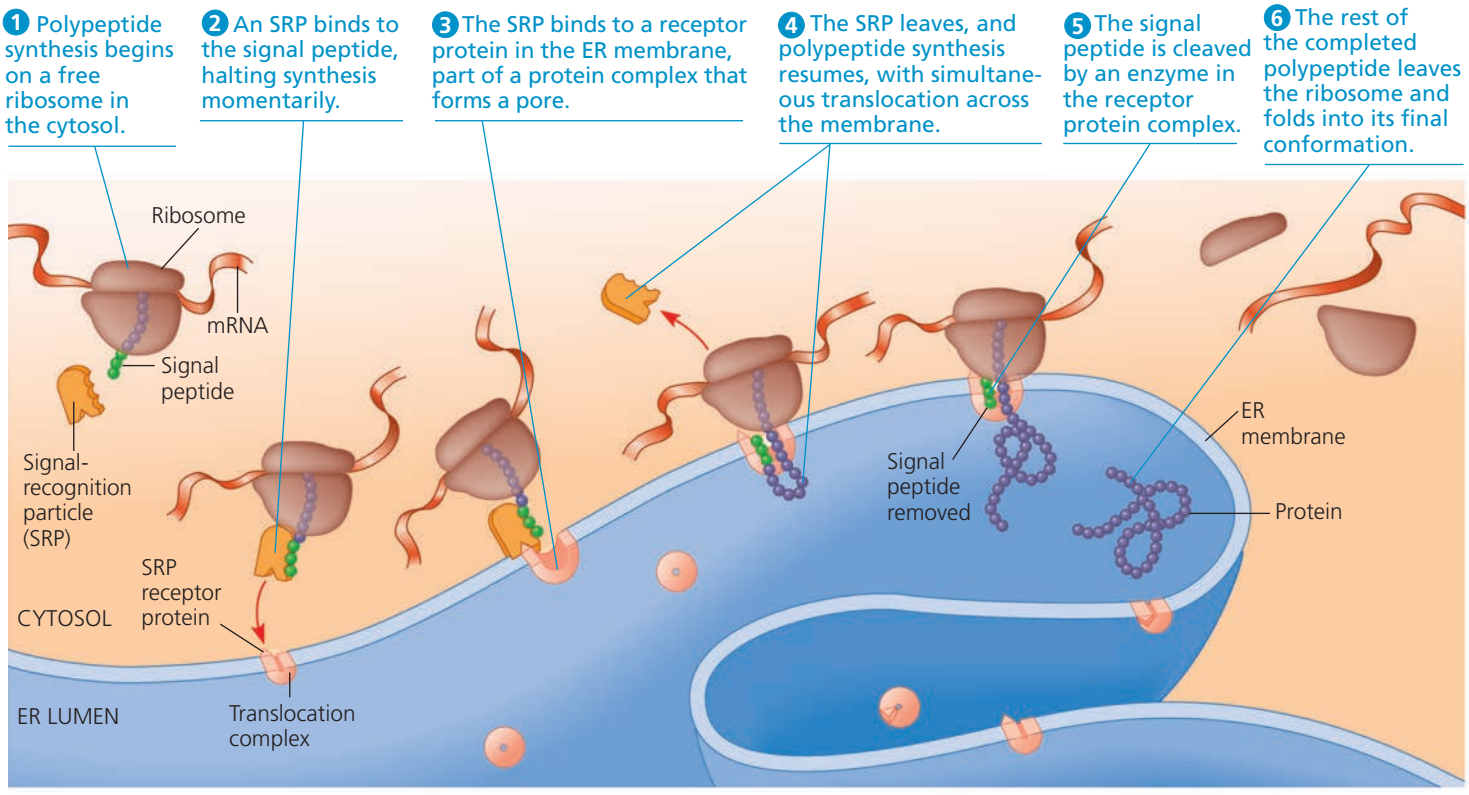
Targeting Polypeptides to Specific Locations

In electron micrographs of eukaryotic cells active in protein synthesis, two populations of ribosomes are evident: free and bound (see Figure 7.10). Free ribosomes are suspended in the cytosol and mostly synthesize proteins that stay in the cytosol and function there. In contrast, bound ribosomes are attached to the cytosolic side of the endoplasmic reticulum (ER) or to the nuclear envelope. Bound ribosomes make proteins of the endomembrane system (see Figure 7.15) as well as proteins secreted from the cell, such as insulin. It is important to note that the ribosomes themselves are identical and can alternate between being free ribosomes one time they are used and being bound ribosomes the next.

What determines whether a ribosome is free in the cytosol or bound to rough ER? Polypeptide synthesis always begins in the cytosol as a free ribosome starts to translate an mRNA molecule. There, the process continues to completion—*unless* the growing polypeptide itself cues the ribosome to attach to the ER. The polypeptides of proteins destined for the endomembrane system or for secretion are marked by a **signal peptide**, which targets the protein to the ER **(Figure 17.22)**. The signal peptide, a sequence of about 20 amino acids at or near the leading end (N-terminus) of the polypeptide, is recognized as it emerges from the ribosome by a protein-RNA complex called a **signal-recognition particle (SRP)**. This particle escorts the ribosome to a receptor protein built into the ER membrane. The receptor is part of a multiprotein translocation complex. Polypeptide synthesis continues there, and the growing polypeptide snakes across the membrane into the ER lumen via a protein pore. The rest of the completed polypeptide, if it is to be secreted from the cell, is released into solution within the ER lumen. Alternatively, if the polypeptide is to be a membrane protein, it remains partially embedded in the ER membrane. In either case, it travels in a transport vesicle to its destination (see, for example, Figure 8.9).

Other kinds of signal peptides are used to target polypeptides to mitochondria, chloroplasts, the interior of the nucleus, and other organelles that are not part of the endomembrane system. The critical difference in these cases is that translation is completed in the cytosol before the polypeptide is imported

Figure 17.22 The signal mechanism for targeting proteins to the ER.



MAKE CONNECTIONS ➤ *If this protein were destined for secretion, what would happen to it after its synthesis was completed? See Figure 8.9.*

Video: Cotranslational Translocation

into the organelle. Translocation mechanisms also vary, but in all cases studied to date, the "postal zip codes" that address proteins for secretion or to cellular locations are signal peptides of some sort. Bacteria also employ signal peptides to target proteins to the plasma membrane for secretion.

Making Multiple Polypeptides in Bacteria and Eukaryotes

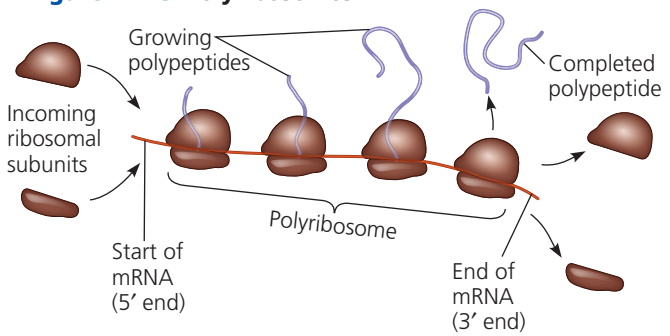
In previous sections, you learned how a single polypeptide is synthesized using the information encoded in an mRNA molecule. When a polypeptide is required in a cell, though, the need is for many copies, not just one.

A single ribosome can make an average-sized polypeptide in less than a minute. In both bacteria and eukaryotes, however, multiple ribosomes translate an mRNA at the same time **(Figure 17.23)**; that is, a single mRNA is used to make many copies of a polypeptide simultaneously. Once a ribosome is far enough past the start codon, a second ribosome can attach to the mRNA, eventually resulting in a number of ribosomes trailing along the mRNA. Such strings of ribosomes, called **polyribosomes** (or **polysomes**), can be seen with an electron microscope; they can be either free or bound. They enable a cell to rapidly make many copies of a polypeptide.

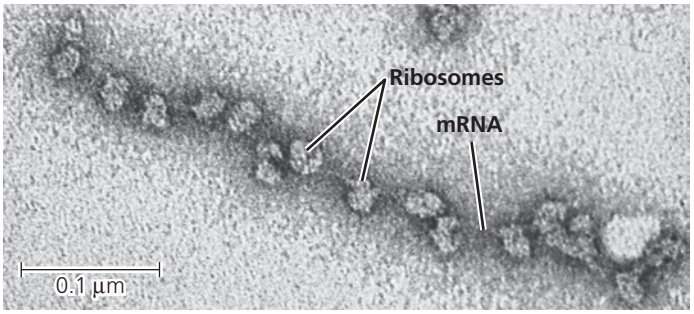
Another way both bacteria and eukaryotes augment the number of copies of a polypeptide is by transcribing multiple mRNAs from the same gene. However, the coordination of the two processes—transcription and translation—differs in the two groups. The most important differences between bacteria and eukaryotes arise from the bacterial cell's lack of compartmental organization. Like a one-room workshop, a bacterial cell ensures a streamlined operation by coupling the two processes. In the absence of a nucleus, it can simultaneously transcribe and translate the same gene **(Figure 17.24)**, and the newly made protein can quickly diffuse to its site of function.

In contrast, the eukaryotic cell's nuclear envelope segregates transcription from translation and provides a compartment for extensive RNA processing. This processing stage includes additional steps, discussed earlier, the regulation of which can help coordinate the eukaryotic cell's elaborate activities. **Figure 17.25** summarizes the path from gene to polypeptide in a eukaryotic cell.

Figure 17.23 Polyribosomes.

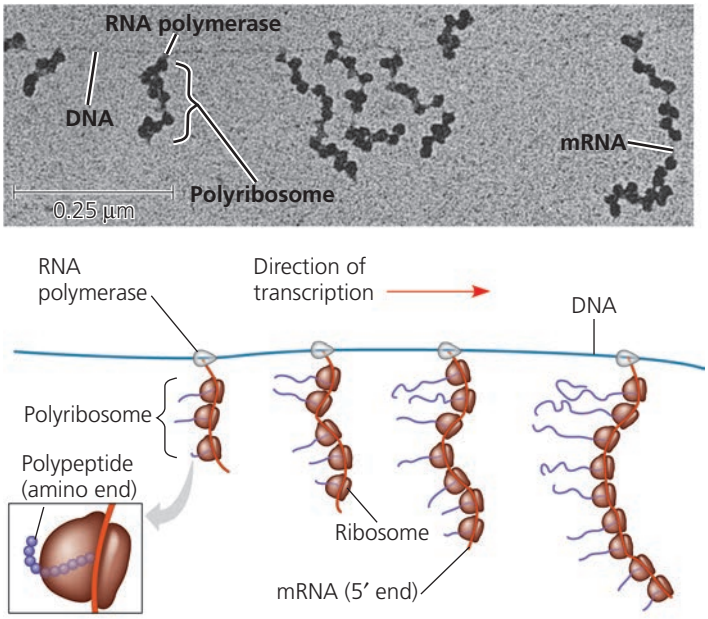


(a) An mRNA molecule is generally translated simultaneously by several ribosomes in clusters called polyribosomes.



(b) This micrograph shows a large polyribosome in a bacterial cell. Growing polypeptides are not visible here (TEM).

Figure 17.24 Coupled transcription and translation in bacteria. In bacterial cells, the translation of mRNA can begin as soon as the leading (5′) end of the mRNA molecule peels away from the DNA template. The micrograph (TEM) shows a strand of *E. coli* DNA being transcribed by RNA polymerase molecules. Attached to each RNA polymerase molecule is a growing strand of mRNA, which is already being translated by ribosomes. The newly synthesized polypeptides are not visible in the micrograph but are shown in the diagram.



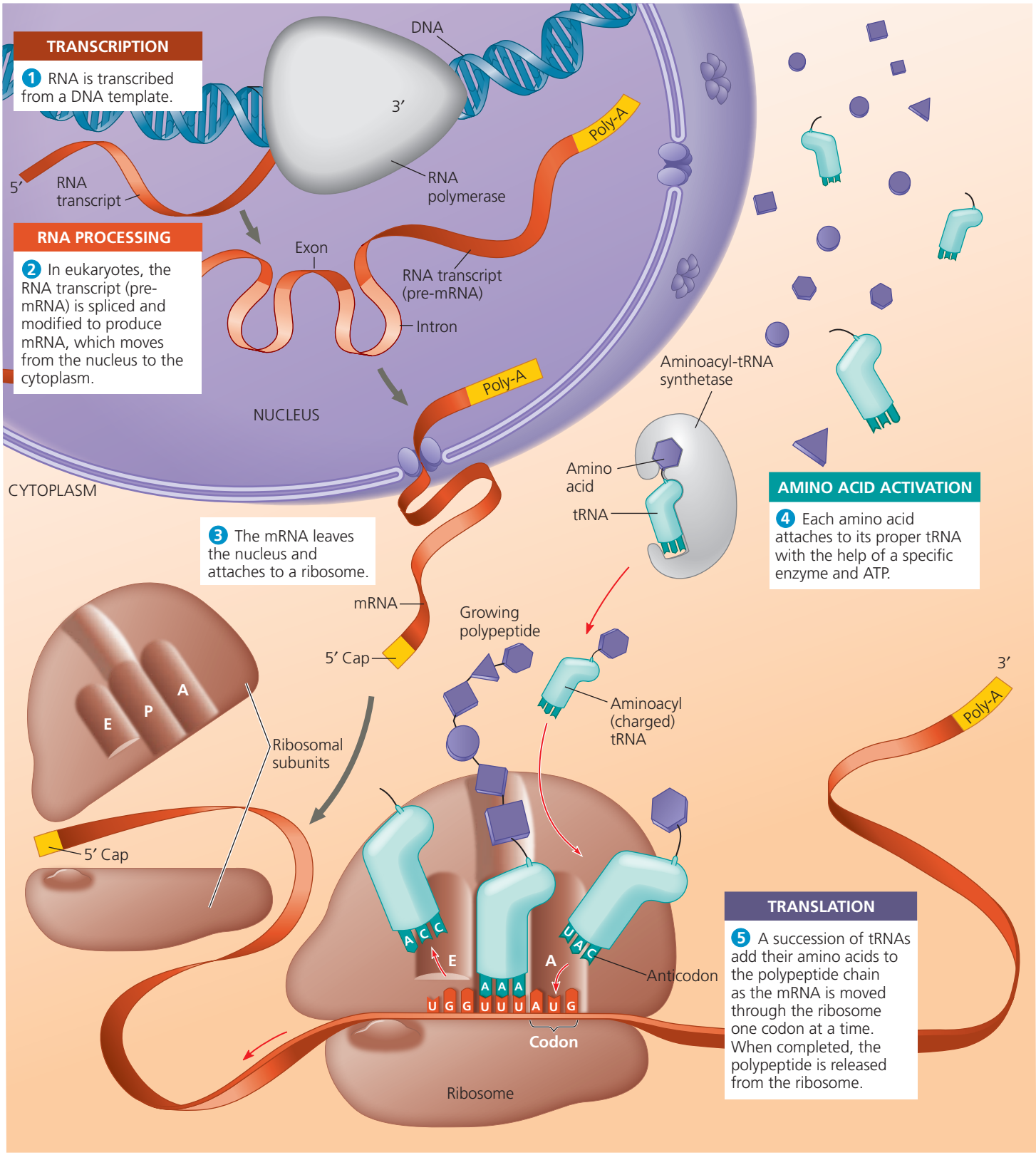
VISUAL SKILLS ➤ *Which one of the mRNA molecules started being transcribed first? On that mRNA, which ribosome started translating the mRNA first?*

CONCEPT CHECK 17.4

1. What two processes ensure that the correct amino acid is added to a growing polypeptide chain?
2. Discuss the different post-translational changes that may be needed to make a functional protein.
3. **WHAT IF? DRAW IT** ➤ Draw a tRNA with the anticodon 3′-CGU-5′. What two different codons could it bind to? Draw each codon on an mRNA, labeling all 5′ and 3′ ends, the tRNA, and the amino acid it carries.
4. **WHAT IF?** ➤ In eukaryotic cells, mRNAs have been found to have a circular arrangement in which proteins hold the poly-A tail near the 5′ cap. How might this increase translation efficiency?

For suggested answers, see Appendix A.

Figure 17.25 A summary of transcription and translation in a eukaryotic cell. This diagram shows the path from one gene to one polypeptide. Each gene in the DNA can be transcribed repeatedly into many identical RNA molecules and each mRNA can be translated repeatedly to yield many identical polypeptide molecules. (Also, remember that the final products of some genes are not polypeptides but RNA molecules that don't get translated, including tRNA and rRNA.) In general, the steps of transcription and translation are similar in bacterial, archaeal, and eukaryotic cells. The major difference is the occurrence of RNA processing in the eukaryotic nucleus. Other significant differences are found in the initiation stages of both transcription and translation and in the termination of transcription. To visualize these processes in their cellular context, see Figure 7.32.



BioFlix® Animation: Protein Synthesis

CONCEPT 17.5

Mutations of one or a few nucleotides can affect protein structure and function

Now that you have explored the process of gene expression, you are ready to understand the effects of changes to the genetic information of a cell. These changes, called **mutations**, are responsible for the huge diversity of genes found among organisms because mutations are the ultimate source of new genes. Earlier, we considered chromosomal rearrangements that affect long segments of DNA (see Figure 15.14); these are considered large-scale mutations. Here we examine small-scale mutations of one or a few nucleotide pairs, including **point mutations**, changes in a single nucleotide pair of a gene.

If a point mutation occurs in a gamete or in a cell that gives rise to gametes, it may be transmitted to offspring and to future generations. If the mutation has an adverse effect on the phenotype of a person, the mutant condition is referred to as a genetic disorder or hereditary disease. For example, we can trace the genetic basis of sickle-cell disease to the mutation of a single nucleotide pair in the gene that encodes the β-globin polypeptide of hemoglobin. The change of a single nucleotide in the DNA's template strand leads to an altered mRNA and the production of an abnormal protein **(Figure 17.26**; also see Figure 5.19). In individuals who are homozygous for the mutant allele, the sickling of red blood cells caused by the altered hemoglobin produces the multiple symptoms associated with sickle-cell disease (see Concept 14.4 and Figure 23.18). Another disorder caused by a point mutation is a heart condition called familial cardiomyopathy, which is responsible for some of the tragic incidents of sudden death in young athletes. Point mutations in several genes encoding muscle proteins have been identified, any of which can lead to this disorder.

Types of Small-Scale Mutations

Let's now consider how small-scale mutations affect proteins. We should first note that many mutations occur in regions outside of protein-coding genes, and any potential effect they have on the phenotype of the organism may be subtle and hard to detect. For this reason, here we'll concentrate on mutations within protein-coding genes. Small-scale mutations within a gene can be divided into two general categories: (1) single nucleotide-pair substitutions and (2) nucleotide-pair insertions or deletions. Insertions and deletions can involve one or more nucleotide pairs.

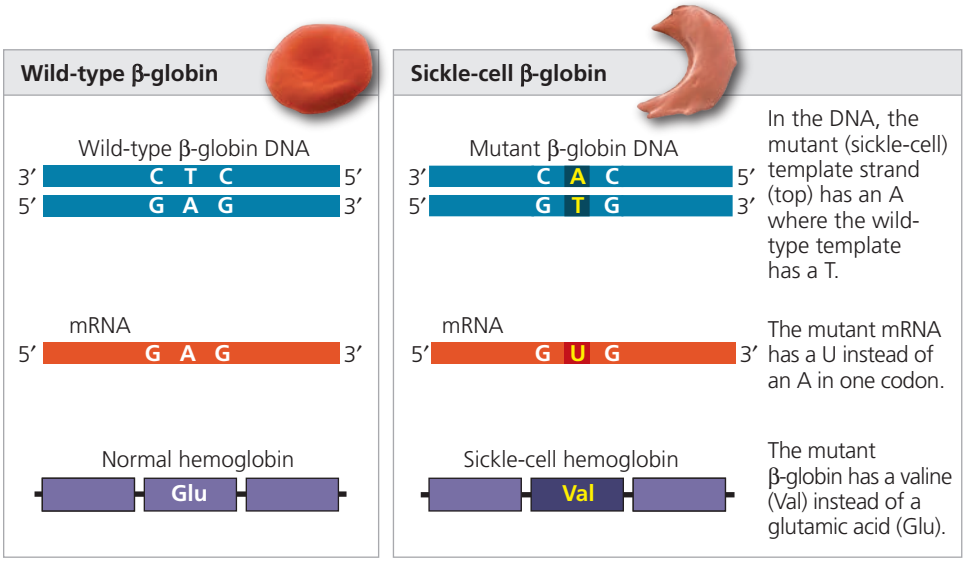
Substitutions

A **nucleotide-pair substitution** is the replacement of one nucleotide and its partner with another pair of nucleotides **(Figure 17.27a)**. Some substitutions have no effect on the encoded protein, owing to the redundancy of the genetic code. For example, if 3′-CCG-5′ on the template strand mutated to 3′-CCA-5′, the mRNA codon that used to be GGC would become GGU, but a glycine would still be inserted at the proper location in the protein (see Figure 17.6). In other words, a change in a nucleotide pair may transform one codon into another that is translated into the same amino acid. Such a change is an example of a **silent mutation**, which has no observable effect on the phenotype. (Silent mutations can occur outside genes as well.) Interestingly, there is evidence that some silent mutations may indirectly affect where or at what level the gene gets expressed, even though the actual protein is the same.

Substitutions that change one amino acid to another one are called **missense mutations**. Such a mutation may have little effect on the protein: The new amino acid may have properties similar to those of the amino acid it replaces, or it may be in a region of the protein where the exact sequence of amino acids is not essential to the protein's function.

However, the nucleotide-pair substitutions of greatest interest are those that cause a major change in a protein. The alteration of a single amino acid in a crucial area of a protein—such as in the part

Figure 17.26 The molecular basis of sickle-cell disease: a point mutation. The allele that causes sickle-cell disease differs from the wild-type (normal) allele by a single DNA nucleotide pair. The micrographs are SEMs of a normal red blood cell (on the left) and a sickled red blood cell (right) from individuals homozygous for wild-type and mutant alleles, respectively.



HHMI Animation: Sickle-Cell Disease
Animation: Mutation Types

Figure 17.27 Types of small-scale mutations that affect mRNA sequence. All but one of the types shown here also affect the amino acid sequence of the encoded polypeptide.

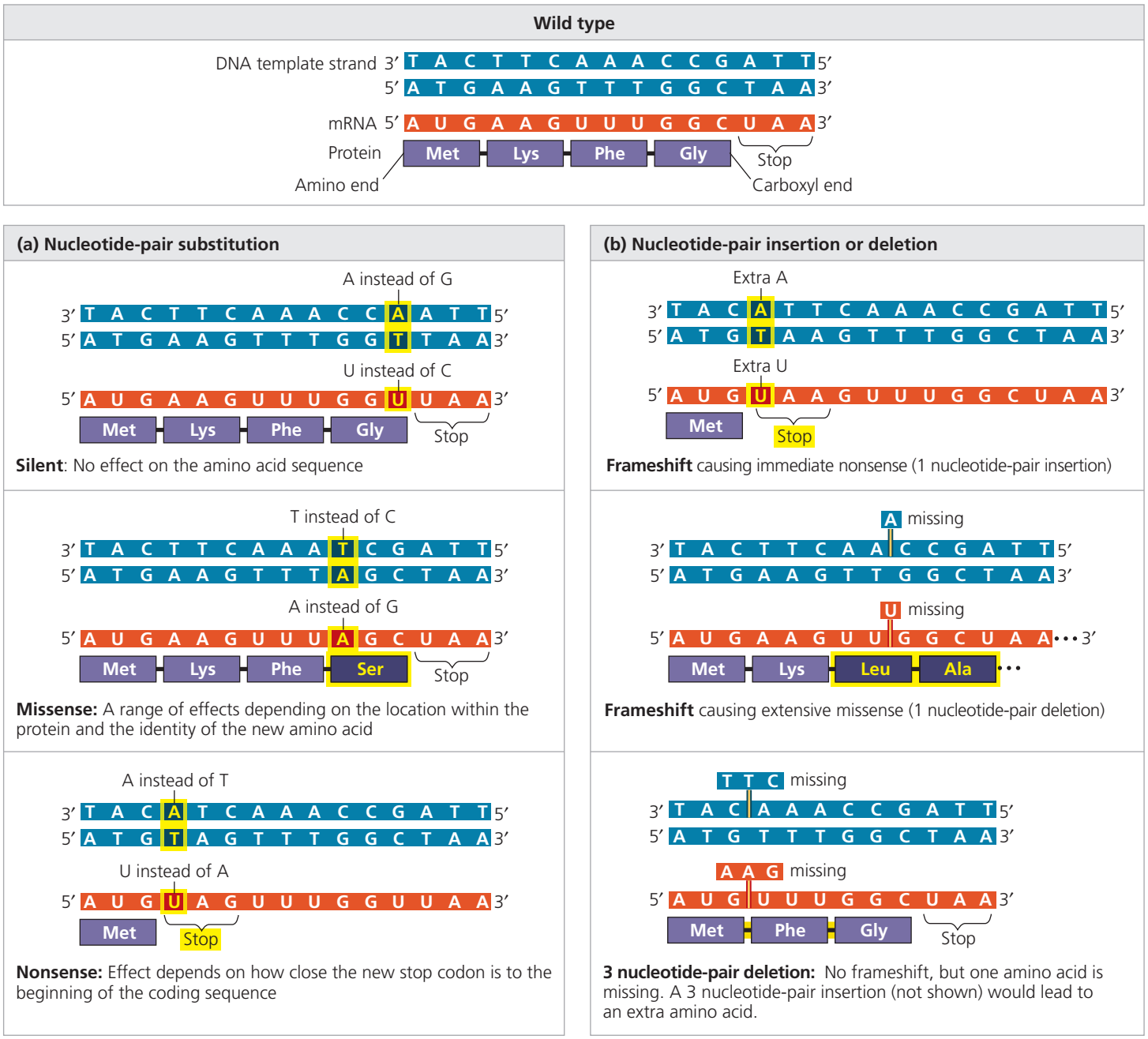


Figure Walkthrough

of the β-globin subunit of hemoglobin shown in Figure 17.26 or in the active site of an enzyme as shown in Figure 6.19—can significantly alter protein activity. Occasionally, such a mutation leads to an improved protein or one with novel capabilities, but much more often such mutations are neutral or detrimental, leading to a useless or less active protein that impairs cellular function.

Substitution mutations are usually missense mutations; that is, the altered codon still codes for an amino acid and thus makes sense, although not necessarily the *right* sense. But a point mutation can also change a codon for an amino acid into a stop codon. This is called a **nonsense mutation**, and it causes translation to be terminated prematurely; the resulting polypeptide will be shorter than the polypeptide encoded by the normal gene. Most nonsense mutations lead to nonfunctional proteins.

In the **Problem-Solving Exercise**, you'll work with a few common single nucleotide-pair substitution mutations in the gene encoding insulin, some or all of which may lead to diabetes. You will classify these mutations into one of

PROBLEM-SOLVING EXERCISE

Are insulin mutations the cause of three infants' neonatal diabetes?

Insulin is a hormone that acts as a key regulator of blood glucose level. In some cases of neonatal diabetes, the gene coding for the insulin protein has a nucleotide-pair substitution mutation that alters the protein structure enough to cause it to malfunction. How can you identify a nucleotide-pair substitution and determine its effect on the amino acid sequence?

Now that it's possible to sequence an individual's whole genome, doctors can use that DNA sequence information to diagnose diseases and identify new treatments. For example, the insulin gene sequence of a patient with neonatal diabetes can be analyzed to determine if it has a mutation and, if so, its effect.

Watch the video in the MasteringBiology Study Area to see how genome sequencing is changing medicine.

MB **ABC News Video: Using Genome Sequencing to Diagnose Gene-Based Diseases**

In this exercise, you will determine the effect of mutations present in a portion of diabetes patients' insulin gene sequences.

Your Approach Suppose you are a medical geneticist presented with three infant patients, all of whom have a nucleotide-pair substitution in their insulin gene. It is your job to analyze each mutation to figure out the effect of the mutation on the amino acid sequence of the insulin protein. To identify the mutation in each patient, you will compare his or her individual insulin complementary DNA (cDNA) sequence to that of the wild-type cDNA. (cDNA is a double-stranded DNA molecule that is based on the mRNA sequence and thus contains only the portion of a gene that is translated—introns are not included. cDNA sequences are commonly used to compare the coding regions of genes.) Identifying the codons that have been changed will tell you which, if any, amino acids are altered in the patient's insulin protein.

Your Data You will analyze the cDNA codons for amino acids 35–54 (of the 110 amino acids) of each patient's insulin protein, so the start codon (AUG) is not present. The sequences of the wild-type cDNA and the patients' cDNA are shown below, arranged in codons.

Wild-type cDNA	5′–CTG GTG GAA GCT CTC TAC CTA GTG TGC GGG GAA CGA GGC TTC TTC TAC ACA CCC AAG ACC–3′
Patient 1 cDNA	5′–CTG GTG GAA GCT CTC TAC CTA GTG TGC GGG GAA CGA GGC TGC TTC TAC ACA CCC AAG ACC–3′
Patient 2 cDNA	5′–CTG GTG GAA GCT CTC TAC CTA GTG TGC GGG GAA CGA GGC TCC TTC TAC ACA CCC AAG ACC–3′
Patient 3 cDNA	5′–CTG GTG GAA GCT CTC TAC CTA GTG TGC GGG GAA CGA GGC TTC TTG TAC ACA CCC AAG ACC–3′

Data from N. Nishi and K. Nanjo, Insulin gene mutations and diabetes, *Journal of Diabetes Investigation* 2:92–100 (2011).

Your Analysis

1. Comparing each patient's cDNA sequence to the wild-type cDNA sequence, circle the codons where a nucleotide-pair substitution mutation has occurred.
2. Use a codon table (see Figure 17.6) to identify the amino acid that will be made by the codon with the mutation in each patient's insulin sequence, and compare it to the amino acid made by the codon in the corresponding wild-type sequence. As is standard practice with DNA sequences, the cDNA *coding* (nontemplate) strand has been provided, so to convert it to mRNA for use with the codon table, you just need to change T to U. Classify each patient's nucleotide-pair substitution mutation: Is it a silent, missense, or nonsense mutation? Explain, for each answer.
3. Compare the structure of the amino acid you identified in each patient's insulin sequence to that of the corresponding amino acid in the wild-type insulin sequence (see Figure 5.14). Given that each patient has neonatal diabetes, discuss how the change of amino acid in each might have affected the insulin protein and thus resulted in the disease.

MB **Instructors**: A version of this Problem-Solving Exercise can be assigned in MasteringBiology. Or a more extensive investigation called "Solve It: Which Insulin Mutations May Result in Disease?" can be assigned.

the types we just described and characterize the change in amino acid sequence.

Insertions and Deletions

Insertions and **deletions** are additions or losses of nucleotide pairs in a gene **(Figure 17.27b)**. These mutations have a disastrous effect on the resulting protein more often than substitutions do. Insertion or deletion of nucleotides may alter the reading frame of the genetic message, the triplet grouping of nucleotides on the mRNA that is read during translation. Such a mutation, called a **frameshift mutation**, occurs whenever the number of nucleotides inserted or deleted is not a multiple of three. All nucleotides downstream of the deletion or insertion will be improperly grouped into codons; the result will be

extensive missense mutations, usually ending sooner or later in a nonsense mutation that leads to premature termination. Unless the frameshift is very near the end of the gene, the protein is almost certain to be nonfunctional. Insertions and deletions also occur outside of coding regions; these are not called frameshift mutations, but can have effects on the phenotype—for instance, they can affect how a gene is expressed.

New Mutations and Mutagens

Mutations can arise in a number of ways. Errors during DNA replication or recombination can lead to nucleotide-pair substitutions, insertions, or deletions, as well as to mutations affecting longer stretches of DNA. If an incorrect nucleotide is added to a growing chain during replication, for example, the base on that nucleotide will then be mismatched with the nucleotide base on the other strand. In many cases, the error will be corrected by DNA proofreading and repair systems (see Concept 16.2). Otherwise, the incorrect base will be used as a template in the next round of replication, resulting in a mutation. Such mutations are called *spontaneous mutations*. It is difficult to calculate the rate at which such mutations occur. Rough estimates have been made of the rate of mutation during DNA replication for both *E. coli* and eukaryotes, and the numbers are similar: About one nucleotide in every 10^{10} is altered, and the change is passed on to the next generation of cells.

A number of physical and chemical agents, called **mutagens**, interact with DNA in ways that cause mutations. In the 1920s, Hermann Muller discovered that X-rays caused genetic changes in fruit flies, and he used X-rays to make *Drosophila* mutants for his genetic studies. But he also recognized an alarming implication of his discovery: X-rays and other forms of high-energy radiation pose hazards to the genetic material of people as well as laboratory organisms. Mutagenic radiation, a physical mutagen, includes ultraviolet (UV) light, which can cause disruptive thymine dimers in DNA (see Figure 16.19).

Chemical mutagens fall into several categories. Nucleotide analogs are chemicals similar to normal DNA nucleotides but that pair incorrectly during DNA replication. Other chemical mutagens interfere with correct DNA replication by inserting themselves into the DNA and distorting the double helix. Still other mutagens cause chemical changes in bases that change their pairing properties.

Researchers have developed a variety of methods to test the mutagenic activity of chemicals. A major application of these tests is the preliminary screening of chemicals to identify those that may cause cancer. This approach makes sense because most carcinogens (cancer-causing chemicals) are mutagenic, and conversely, most mutagens are carcinogenic.

What Is a Gene? *Revisiting the Question*

Our definition of a gene has evolved over the past few chapters, as it has through the history of genetics. We began with the Mendelian concept of a gene as a discrete unit of inheritance that affects a phenotypic character (Chapter 14). We saw that Morgan and his colleagues assigned such genes to specific loci on chromosomes (Chapter 15). We went on to view a gene as a region of specific nucleotide sequence along the length of the DNA molecule of a chromosome (Chapter 16). Finally, in this chapter, we have considered a functional definition of a gene as a DNA sequence that codes for a specific polypeptide chain or a functional RNA molecule, such as a tRNA. All these definitions are useful, depending on the context in which genes are being studied.

We have noted that merely saying a gene codes for a polypeptide is an oversimplification. Most eukaryotic genes contain noncoding segments (such as introns), so large portions of these genes have no corresponding segments in polypeptides. Molecular biologists also often include promoters and certain other regulatory regions of DNA within the boundaries of a gene. These DNA sequences are not transcribed, but they can be considered part of the functional gene because they must be present for transcription to occur. Our definition of a gene must also be broad enough to include the DNA that is transcribed into rRNA, tRNA, and other RNAs that are not translated. These genes have no polypeptide products but play crucial roles in the cell. Thus, we arrive at the following definition: *A gene is a region of DNA that can be expressed to produce a final functional product that is either a polypeptide or an RNA molecule*.

When considering phenotypes, however, it is often useful to start by focusing on genes that code for polypeptides. In this chapter, you have learned in molecular terms how a typical gene is expressed—by transcription into RNA and then translation into a polypeptide that forms a protein of specific structure and function. Proteins, in turn, bring about an organism's observable phenotype.

A given type of cell expresses only a subset of its genes. This is an essential feature in multicellular organisms: You'd be in trouble if the lens cells in your eyes started expressing the genes for hair proteins, which are normally expressed only in hair follicle cells! Gene expression is precisely regulated, which we'll explore in the next chapter, beginning with the simpler case of bacteria and continuing with eukaryotes.

CONCEPT CHECK 17.5

1. What happens when one nucleotide pair is lost from the middle of the coding sequence of a gene?
2. **MAKE CONNECTIONS ➤** Individuals heterozygous for the sickle-cell allele are generally healthy but show phenotypic effects of the allele under some circumstances (see Figure 14.17). Explain in terms of gene expression.
3. **WHAT IF? DRAW IT ➤** The template strand of a gene includes this sequence:
   3′-TACTTGTCCGATATC-5′. It is mutated to
   3′-TACTTGTCCAATATC-5′. For both wild-type and mutant sequences, draw the double-stranded DNA, the resulting mRNA, and the amino acid sequence each encodes. What is the effect of the mutation on the amino acid sequence?

For suggested answers, see Appendix A.

17 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 17.1

VOCAB SELF-QUIZ goo.gl/Rn5Uax

Genes specify proteins via transcription and translation *(pp. 386–392)*

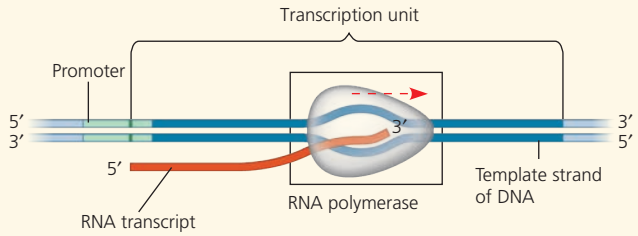
- Beadle and Tatum's studies of mutant strains of *Neurospora* led to the one gene–one polypeptide hypothesis. During **gene expression**, the information encoded in genes is used to make specific polypeptide chains (enzymes and other proteins) or RNA molecules.
- **Transcription** is the synthesis of RNA complementary to a **template strand** of DNA. **Translation** is the synthesis of a polypeptide whose amino acid sequence is specified by the nucleotide sequence in **messenger RNA (mRNA)**.
- Genetic information is encoded as a sequence of nonoverlapping nucleotide triplets, or **codons**. A codon in mRNA either is translated into an amino acid (61 of the 64 codons) or serves as a stop signal (3 codons). Codons must be read in the correct **reading frame**.

? *Describe the process of gene expression, by which a gene affects the phenotype of an organism.*

CONCEPT 17.2

Transcription is the DNA-directed synthesis of RNA: *a closer look* *(pp. 392–394)*

- RNA synthesis is catalyzed by **RNA polymerase**, which links together RNA nucleotides complementary to a DNA template strand. This process follows the same base-pairing rules as DNA replication, except that in RNA, uracil substitutes for thymine.



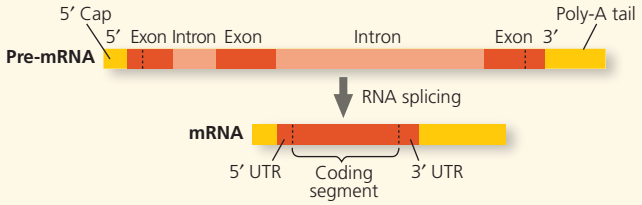
- The three stages of transcription are initiation, elongation, and termination. A **promoter**, often including a **TATA box** in eukaryotes, establishes where RNA synthesis is initiated. **Transcription factors** help eukaryotic RNA polymerase recognize promoter sequences, forming a **transcription initiation complex**. Termination differs in bacteria and eukaryotes.

? *What are the similarities and differences in the initiation of gene transcription in bacteria and eukaryotes?*

CONCEPT 17.3

Eukaryotic cells modify RNA after transcription *(pp. 395–397)*

- Eukaryotic mRNAs undergo **RNA processing**, which includes RNA splicing, the addition of a modified nucleotide **5′ cap** to the 5′ end, and the addition of a **poly-A tail** to the 3′ end. The processed mRNA includes an untranslated region (5′ UTR or 3′ UTR) at each end of the coding segment.
- Most eukaryotic genes are split into segments: They have **introns** interspersed among the **exons** (the regions included in the mRNA). In **RNA splicing**, introns are removed and exons joined. RNA splicing is typically carried out by **spliceosomes**, but in some cases, RNA alone catalyzes its own splicing. The catalytic ability of some RNA molecules, called **ribozymes**, derives from the inherent properties of RNA. The presence of introns allows for **alternative RNA splicing**.

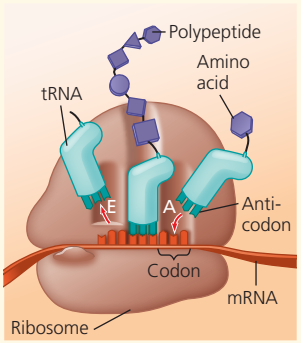


? *What function do the 5′ cap and the poly-A tail serve on a eukaryotic mRNA?*

CONCEPT 17.4

Translation is the RNA-directed synthesis of a polypeptide: *a closer look* *(pp. 397–406)*

- A cell translates an mRNA message into protein using **transfer RNAs (tRNAs)**. After being bound to a specific amino acid by an **aminoacyl-tRNA synthetase**, a tRNA lines up via its **anticodon** at the complementary codon on mRNA. A **ribosome**, made up of **ribosomal RNAs (rRNAs)** and proteins, facilitates this coupling with binding sites for mRNA and tRNA.
- Ribosomes coordinate the three stages of translation: initiation, elongation, and termination. The formation of peptide bonds between amino acids is catalyzed by rRNAs as tRNAs move through the **A** and **P sites** and exit through the **E site**.



- After translation, during protein processing, proteins may be modified by cleavage or by attachment of sugars, lipids, phosphates, or other chemical groups.
- Free ribosomes in the cytosol initiate synthesis of all proteins, but proteins with a **signal peptide** are synthesized on the ER.
- A gene can be transcribed by multiple RNA polymerases simultaneously. Also, a single mRNA molecule can be translated simultaneously by a number of ribosomes, forming a **polyribosome**. In bacteria, these processes are coupled, but in eukaryotes they are separated in space and time by the nuclear membrane.

? *Describe how tRNAs function in the context of the ribosome in building a polypeptide.*

CONCEPT 17.5

Mutations of one or a few nucleotides can affect protein structure and function *(pp. 407–410)*

- Small-scale **mutations** include **point mutations**, changes in one DNA nucleotide pair, which may lead to production of nonfunctional proteins. **Nucleotide-pair substitutions** can cause **missense** or **nonsense mutations**. Nucleotide-pair **insertions** or **deletions** may produce **frameshift mutations**.
- Spontaneous mutations can occur during DNA replication and recombination. Chemical and physical **mutagens** cause DNA damage that can alter genes.

? *What will be the results of chemically modifying one nucleotide base of a gene? What role is played by DNA repair systems in the cell?*

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–7 can be found in the Study Area in MasteringBiology.

8. Would the coupling of the processes shown in Figure 17.24 be found in a eukaryotic cell? Explain why or why not.

9. Complete the following table:

PRACTICE TEST goo.gl/iAsVgL

Type of RNA	Functions
Messenger RNA (mRNA)	
Transfer RNA (tRNA)	
	In a ribosome, plays a structural role; as a ribozyme, plays a catalytic role (catalyzes peptide bond formation)
Primary transcript	
Small RNAs in the spliceosome	

10. **EVOLUTION CONNECTION** Most amino acids are coded for by a set of similar codons (see Figure 17.6). Propose at least one evolutionary explanation to account for this pattern.

11. **SCIENTIFIC INQUIRY** During the sequence analysis of a eukaryotic gene, you discover that it contains a restriction site for the enzyme NotI. This site is 1500 bp away from the restriction site for the enzyme XbaI. When the cDNA for the gene is treated with both the enzymes, the two sites are found to be only 110 bp apart. What is the most likely reason for this result?

12. **WRITE ABOUT A THEME: INFORMATION** Evolution accounts for the unity and diversity of life, and the continuity of life is based on heritable information in the form of DNA. In a short essay (100–150 words), discuss how the fidelity with which DNA is inherited is related to the processes of evolution. (Review the discussion of proofreading and DNA repair in Concept 16.2.)

13. **SYNTHESIZE YOUR KNOWLEDGE**

Some mutations result in proteins that function well at one temperature but are nonfunctional at a different (usually higher) temperature. Siamese cats have such a "temperature-sensitive" mutation in a gene encoding an enzyme that makes dark pigment in the fur. The mutation results in the breed's distinctive point markings and lighter body color (see the photo). Using this information and what you learned in the chapter, explain the pattern of the cat's fur pigmentation.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

18 Control of Gene Expression

Figure 18.1 How can this fish's eyes see equally well in both air and water?

KEY CONCEPTS

Beauty in the Eye of the Beholder

The fish in **Figure 18.1** is keeping an eye out for predators—or, more precisely, both halves of each eye! *Anableps anableps*, commonly known as "cuatro ojos" ("four eyes"), glides through freshwater lakes and ponds in Central and South America with the upper half of each eye protruding from the water. The eye's upper half is particularly well-suited for aerial vision and the lower half for aquatic vision. The molecular basis of this specialization has recently been revealed: The cells of the two parts of the eye express a slightly different set of genes involved in vision, even though these two groups of cells are quite similar and contain identical genomes. What is the biological mechanism underlying the difference in gene expression that makes this remarkable feat possible?

A hallmark of prokaryotic and eukaryotic cells alike—from a bacterium to the cells of a fish—is their intricate and precise regulation of gene expression. In this chapter, we first explore how bacteria regulate expression of their genes in response to different environmental conditions. We then examine the general mechanisms by which eukaryotes regulate gene expression, including the many roles played by RNA molecules. In the final two sections, we explore the role of gene regulation in both embryonic development, as the ultimate example of proper gene regulation, and cancer, as an illustration of what happens when regulation goes awry. Orchestrating proper gene expression by all cells is crucial to the functions of life.

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

CONCEPT 18.1

Bacteria often respond to environmental change by regulating transcription

Bacterial cells that can conserve resources and energy have a selective advantage over cells that are unable to do so. Thus, natural selection has favored bacteria that express only the genes whose products are needed by the cell.

Consider, for instance, an individual *Escherichia coli* (*E. coli*) cell living in the erratic environment of a human colon, dependent for its nutrients on the whimsical eating habits of its host. If the environment is lacking in the amino acid tryptophan, which the bacterium needs to survive, the cell responds by activating a metabolic pathway that makes tryptophan from another compound. If the human host later eats a tryptophan-rich meal, the bacterial cell stops producing tryptophan, thus avoiding wasting resources to produce a substance that is readily available from the surrounding solution.

A metabolic pathway can be controlled on two levels, as shown for the synthesis of tryptophan in **Figure 18.2**. First, cells can adjust the activity of enzymes already present. This is a fairly rapid response, which relies on the sensitivity of many enzymes to chemical cues that increase or decrease their catalytic activity (see Concept 6.5). The activity of the first enzyme in the pathway is inhibited by the pathway's end product—tryptophan, in this case **(Figure 18.2a)**. Thus, if tryptophan accumulates in a cell, it shuts down the synthesis of more tryptophan by inhibiting enzyme activity. Such *feedback inhibition*, typical of anabolic (biosynthetic) pathways, allows a cell to adapt to short-term fluctuations in the supply of a substance it needs (see Figure 6.21).

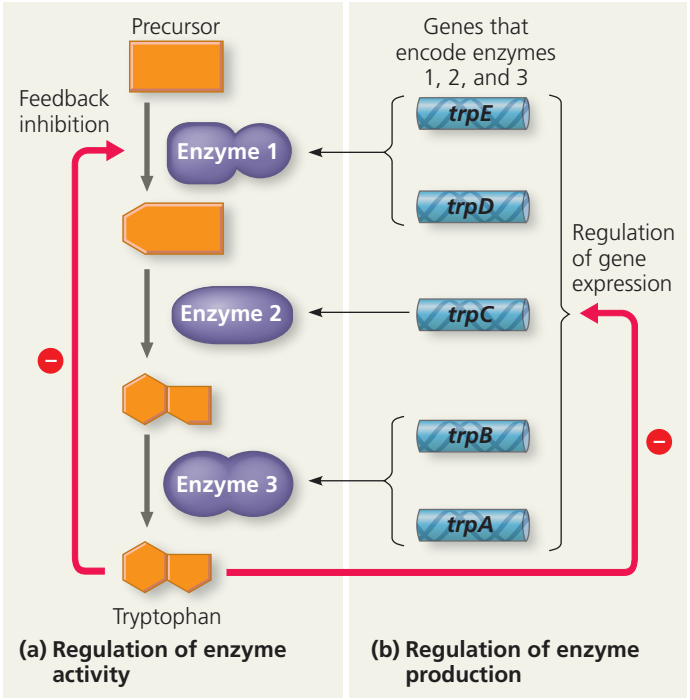
Second, cells can adjust the production level of certain enzymes; that is, they can regulate the expression of the genes encoding the enzymes. If, in our example, the environment provides all the tryptophan the cell needs, the cell stops making the enzymes that catalyze the synthesis of tryptophan **(Figure 18.2b)**. In this case, the control of enzyme production occurs at the level of transcription, the synthesis of messenger RNA from the genes that code for these enzymes.

Regulation of the tryptophan synthesis pathway is just one example of how bacteria tune their metabolism to changing environments. Many genes of the bacterial genome are switched on or off by changes in the metabolic status of the cell. One basic mechanism for this control of gene expression in bacteria, described as the *operon model*, was discovered in 1961 by François Jacob and Jacques Monod at the Pasteur Institute in Paris. Let's see what an operon is and how it works.

Operons: The Basic Concept

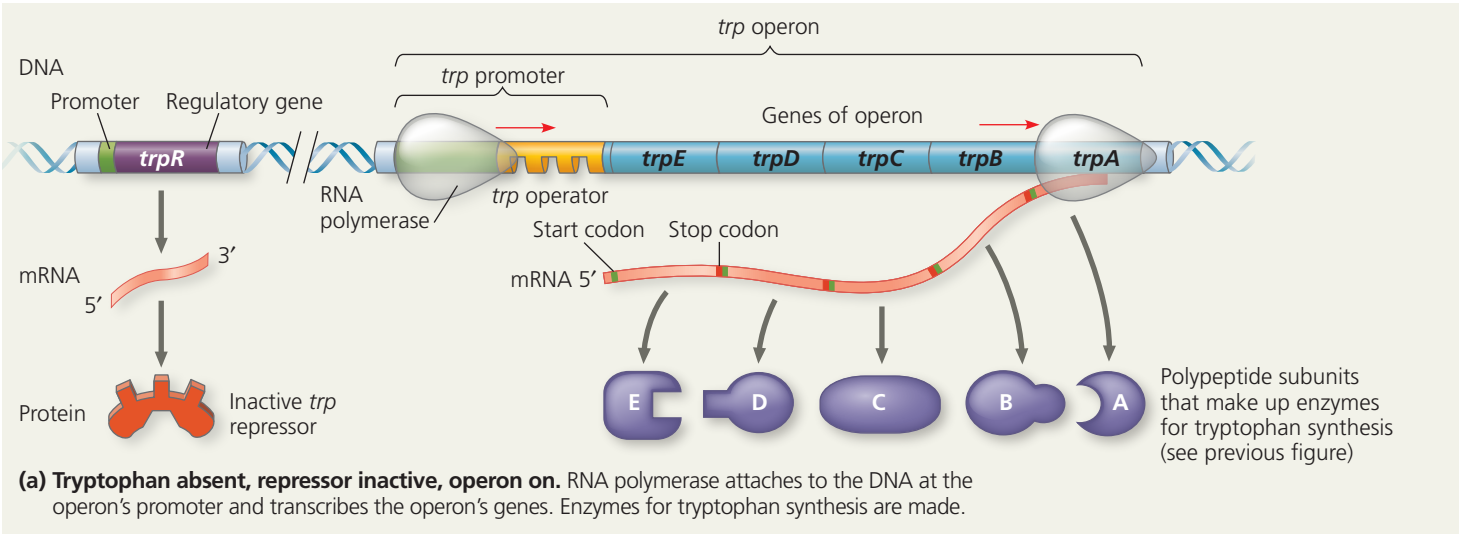
E. coli synthesizes the amino acid tryptophan from a precursor molecule in the three-step pathway shown in Figure 18.2.

Figure 18.2 Regulation of a metabolic pathway. In the pathway for tryptophan synthesis, an abundance of tryptophan can both **(a)** inhibit the activity of the first enzyme in the pathway (feedback inhibition), a rapid response, and **(b)** repress expression of the genes encoding all subunits of the enzymes in the pathway, a longer-term response. Genes *trpE* and *trpD* encode the two subunits of enzyme 1, and genes *trpB* and *trpA* encode the two subunits of enzyme 3. (The genes were named before the order in which they functioned in the pathway was determined.) The symbol ⊖ stands for inhibition.

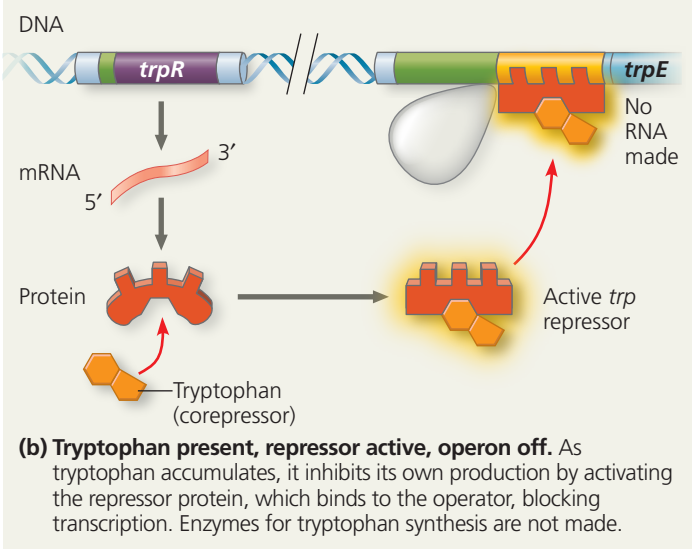


Each reaction in the pathway is catalyzed by a specific enzyme, and the five genes that code for the subunits of these enzymes are clustered together on the bacterial chromosome. A single promoter serves all five genes, which together constitute a transcription unit. (Recall that a promoter is a site where RNA polymerase can bind to DNA and begin transcription; see Figure 17.8.) Thus, transcription gives rise to one long mRNA molecule that codes for the five polypeptides making up the enzymes in the tryptophan pathway **(Figure 18.3a)**. The cell can translate this one mRNA into five separate polypeptides because the mRNA is punctuated with start and stop codons that signal where the coding sequence for each polypeptide begins and ends.

A key advantage of grouping genes of related function into one transcription unit is that a single "on-off switch" can control the whole cluster of functionally related genes; in other words, these genes are *coordinately controlled*. When an *E. coli* cell must make tryptophan for itself because its surrounding environment lacks this amino acid, all the enzymes for the metabolic pathway are synthesized at the same time. The on-off switch is a segment of DNA called an **operator**. Both its location and name suit its function: Positioned within the promoter or, in some cases, between the promoter and the enzyme-coding genes, the operator controls the access



(a) Tryptophan absent, repressor inactive, operon on. RNA polymerase attaches to the DNA at the operon's promoter and transcribes the operon's genes. Enzymes for tryptophan synthesis are made.



(b) Tryptophan present, repressor active, operon off. As tryptophan accumulates, it inhibits its own production by activating the repressor protein, which binds to the operator, blocking transcription. Enzymes for tryptophan synthesis are not made.

Figure 18.3 The *trp* operon in *E. coli*: regulated synthesis of repressible enzymes. Tryptophan is an amino acid produced by an anabolic pathway catalyzed by three enzymes (see Figure 18.2). **(a)** The five genes encoding the polypeptide subunits of the enzymes in this pathway are grouped, along with a promoter, into the *trp* operon. The *trp* operator (the repressor binding site) is located within the *trp* promoter (the RNA polymerase binding site). **(b)** Accumulation of tryptophan, the end product of the pathway, represses transcription of the *trp* operon, thus blocking synthesis of all the enzymes in the pathway and shutting down tryptophan production.

VISUAL SKILLS ➤ *Describe what happens to the* trp *operon as the cell uses up its store of tryptophan.*

of RNA polymerase to the genes. Together, the operator, the promoter, and the genes they control—the entire stretch of DNA required for enzyme production for the tryptophan pathway—constitute an **operon**. The *trp* operon (*trp* for tryptophan) is one of many operons in the *E. coli* genome (see Figure 18.3a).

If the operator is the operon's switch for controlling transcription, how does this switch work? By itself, the *trp* operon is turned on; that is, RNA polymerase can bind to the promoter and transcribe the genes of the operon. The *trp* operon can be switched off by a protein that is called the *trp* repressor. A **repressor** binds to the operator and blocks attachment of RNA polymerase to the promoter, preventing transcription of the genes **(Figure 18.3b)**. A repressor protein is specific for the operator of a particular operon. For example, the *trp* repressor, which switches off the *trp* operon by binding to the *trp* operator, has no effect on other operons in the *E. coli* genome.

A repressor protein is encoded by a **regulatory gene**—in this case, a gene called *trpR*; *trpR* is located some distance from the *trp* operon and has its own promoter. Regulatory genes are expressed continuously, although at a low rate, and a few *trp* repressor molecules are always present in *E. coli* cells. Why, then, is the *trp* operon not switched off permanently? First, the binding of repressors to operators is reversible. An operator alternates between two states: one with the repressor bound and one without the repressor bound. The relative duration of the repressor-bound state is higher when more active repressor molecules are present. Second, the *trp* repressor, like most regulatory proteins, is an allosteric protein, with two alternative shapes: active and inactive (see Figure 6.20). The *trp* repressor is synthesized in the inactive form, which has little affinity for the *trp* operator. Only when a tryptophan molecule binds to the *trp* repressor at an allosteric site does the repressor protein change to the active form that can attach to the operator, turning the operon off.

Tryptophan functions in this system as a **corepressor**, a small molecule that cooperates with a repressor protein to switch an operon off. As tryptophan accumulates, more tryptophan molecules associate with *trp* repressor molecules, which can then bind to the *trp* operator and shut down production of the tryptophan pathway enzymes. If the cell's tryptophan level drops, many fewer *trp* repressor proteins would have tryptophan bound, rendering them inactive; they would dissociate from the operator, allowing transcription of the operon's genes to resume. The *trp* operon is one example of how gene expression can respond to changes in the cell's internal and external environment.

Repressible and Inducible Operons: Two Types of Negative Gene Regulation

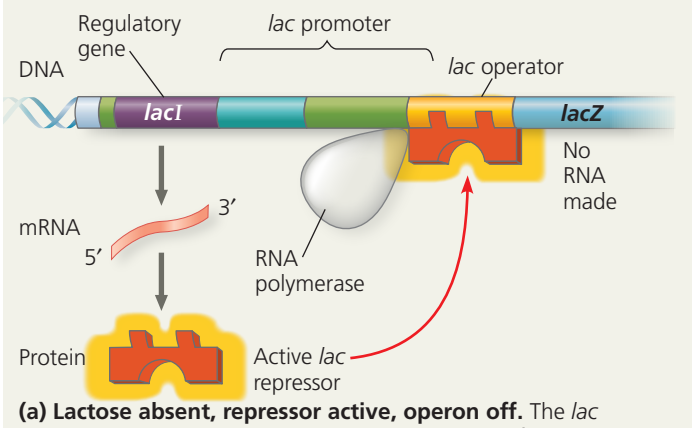
The *trp* operon is said to be a *repressible operon* because its transcription is usually on but can be inhibited (repressed) when a specific small molecule (in this case, tryptophan) binds allosterically to a regulatory protein. In contrast, an *inducible operon* is usually off but can be stimulated (induced) to be on when a specific small molecule interacts with a different regulatory protein. The classic example of an inducible operon is the *lac* operon (*lac* stands for "lactose").

The disaccharide lactose (milk sugar) is available to *E. coli* in the human colon if the host drinks milk or eats a dairy product. Lactose metabolism begins with hydrolysis of the disaccharide into its component monosaccharides (glucose and galactose), a reaction catalyzed by the enzyme β-galactosidase. Only a few molecules of this enzyme are present in an *E. coli* cell growing in the absence of lactose. If lactose is added to the bacterium's environment, however, the number of β-galactosidase molecules in the cell increases 1,000-fold within about 15 minutes. How can a cell ramp up enzyme production this quickly?

The gene for β-galactosidase (*lacZ*) is part of the *lac* operon, which includes two other genes coding for enzymes that function in the use of lactose **(Figure 18.4)**. The entire transcription unit is under the command of one main operator and promoter. The regulatory gene, *lacI*, located outside the *lac* operon, codes for an allosteric repressor protein that can switch off the *lac* operon by binding to the *lac* operator. So far, this sounds just like regulation of the *trp* operon, but there is one important difference. Recall that the *trp* repressor protein is inactive by itself and requires tryptophan as a corepressor in order to bind to the operator. The *lac* repressor, in contrast, is active by itself, binding to the operator and switching the *lac* operon off. In this case, a specific small molecule, called an **inducer**, *inactivates* the repressor.

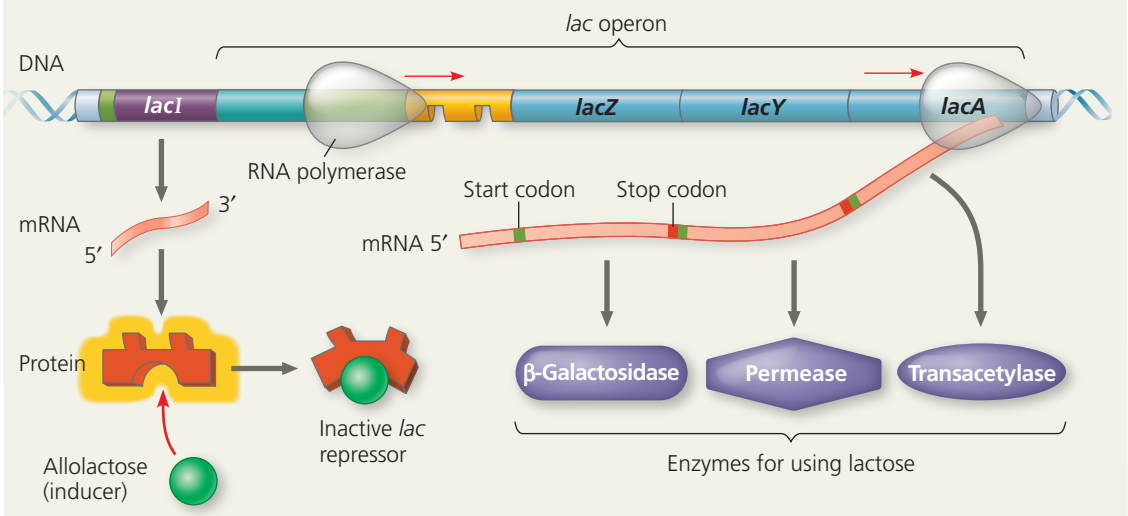
For the *lac* operon, the inducer is allolactose, an isomer of lactose formed in small amounts from lactose that enters the cell. In the absence of lactose (and hence allolactose), the *lac* repressor is in its active shape and binds to the operator; thus, the genes of the *lac* operon are silenced **(Figure 18.4a)**. If lactose is added to the cell's surroundings, allolactose binds to the *lac* repressor and alters its shape so the repressor can no longer bind to the operator. Without the *lac* repressor bound, the *lac* operon is transcribed into mRNA, and the enzymes for using lactose are made **(Figure 18.4b)**.

In the context of gene regulation, the enzymes of the lactose pathway are referred to as *inducible enzymes* because their synthesis is induced by a chemical signal (allolactose, in this case). Analogously, the enzymes for tryptophan synthesis are said to be repressible. *Repressible enzymes* generally function in anabolic pathways, which synthesize essential end products from raw materials (precursors). By suspending production



(a) Lactose absent, repressor active, operon off. The *lac* repressor is innately active, and in the absence of lactose it switches off the operon by binding to the operator. The enzymes for using lactose are not made.

Animation: The *lac* Operon in *E. coli*



(b) Lactose present, repressor inactive, operon on. Allolactose, an isomer of lactose, binds to the repressor, inactivating it and "derepressing" the operon. The inactive repressor cannot bind to the operator, and so the genes of the *lac* operon are transcribed, and the enzymes for using lactose are made.

Figure 18.4 The *lac* operon in *E. coli*: regulated synthesis of inducible enzymes. *E. coli* uses three enzymes to take up and metabolize lactose, the genes for which are clustered in the *lac* operon. The first gene, *lacZ*, codes for β-galactosidase, which hydrolyzes lactose to glucose and galactose. The second, *lacY*, codes for a permease, the membrane protein that transports lactose into the cell. The third, *lacA*, codes for transacetylase, an enzyme that detoxifies other molecules entering the cell via the permease. Unusually, the gene for the *lac* repressor, *lacI*, is adjacent to the *lac* operon. The function of the teal region within the promoter will be revealed in Figure 18.5.

of an end product when it is already present in sufficient quantity, the cell can allocate its organic precursors and energy for other uses. In contrast, inducible enzymes usually function in catabolic pathways, which break down a nutrient to simpler molecules. By producing the appropriate enzymes only when the nutrient is available, the cell avoids wasting energy and precursors making proteins that are not needed.

Regulation of both the *trp* and *lac* operons involves the *negative* control of genes because the operons are switched off by the active form of their respective repressor protein. It may be easier to see this for the *trp* operon, but it is also true for the *lac* operon. In the case of the *lac* operon, allolactose induces enzyme synthesis not by directly activating the *lac* operon, but by freeing it from the negative effect of the repressor (see Figure 18.4b). Gene regulation is said to be *positive* only when a regulatory protein interacts directly with the genome to switch transcription on.

Positive Gene Regulation

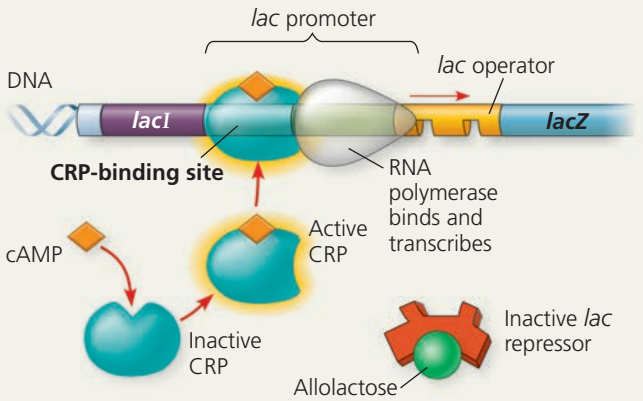
When glucose and lactose are both present in its environment, *E. coli* preferentially uses glucose. The enzymes for glucose breakdown in glycolysis (see Figure 10.9) are continually present. Only when lactose is present *and* glucose is in short supply does *E. coli* use lactose as an energy source, and only then does it synthesize appreciable quantities of the enzymes for lactose breakdown.

How does the *E. coli* cell sense the glucose concentration and relay this information to the *lac* operon? Again, the mechanism depends on the interaction of an allosteric regulatory protein with a small organic molecule, **cyclic AMP (cAMP)** in this case, which accumulates when glucose is scarce (see Figure 9.11 for the structure of cAMP). The regulatory protein, called *cAMP receptor protein* (*CRP*), is an **activator**, a protein that binds to DNA and stimulates transcription of a gene. When cAMP binds to this regulatory protein, CRP assumes its active shape and can attach to a specific site at the upstream end of the *lac* promoter **(Figure 18.5a)**. This attachment increases the affinity of RNA polymerase for the *lac* promoter, which is actually rather low even when no *lac* repressor is bound to the operator. By facilitating the binding of RNA polymerase to the promoter and thereby increasing the rate of transcription of the *lac* operon, the attachment of CRP to the promoter directly stimulates gene expression. Therefore, this mechanism qualifies as positive regulation.

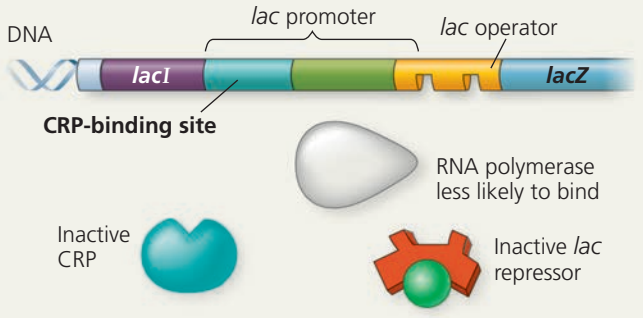
If the amount of glucose in the cell increases, the cAMP concentration falls, and without cAMP, CRP detaches from the *lac* operon. Because CRP is inactive, RNA polymerase binds less efficiently to the promoter, and transcription of the *lac* operon proceeds only at a low level, even when lactose is present **(Figure 18.5b)**. Thus, the *lac* operon is under dual control: negative control by the *lac* repressor and positive control by CRP. The state of the *lac* repressor (with allolactose bound or without it) determines whether or not transcription of the *lac* operon's genes occurs at all; the state of CRP (with bound cAMP or without it) controls the *rate* of transcription if the operon is repressor-free. It is as though the operon has both an on-off switch and a volume control.

In addition to regulating the *lac* operon, CRP helps regulate other operons that encode enzymes used in catabolic pathways. All told, it may affect the expression of more than 100 genes in *E. coli*. When glucose is plentiful and CRP is inactive, the synthesis of enzymes that catabolize compounds other than glucose generally slows down. The ability to catabolize other compounds, such as lactose, enables a cell deprived of glucose to survive. The compounds present in any given cell at the moment determine which operons are switched on—the result of simple interactions of activator and repressor proteins with the promoters of the genes in question.

Figure 18.5 Positive control of the lac operon by cAMP receptor protein (CRP). RNA polymerase has high affinity for the *lac* promoter only when CRP is bound to a DNA site at the upstream end of the promoter. CRP, in turn, attaches to its DNA site only when associated with cyclic AMP (cAMP), whose concentration in the cell rises when the glucose concentration falls. Thus, when glucose is present, even if lactose is also available, the cell preferentially catabolizes glucose and makes very little of the enzymes for using lactose.



(a) Lactose present, glucose scarce (cAMP level high): abundant *lac* mRNA synthesized. If glucose is scarce, the high level of cAMP activates CRP, which binds to the promoter and increases RNA polymerase binding there. The *lac* operon produces large amounts of mRNA coding for the enzymes that the cell needs for use of lactose.



(b) Lactose present, glucose present (cAMP level low): little *lac* mRNA synthesized. When glucose is present, cAMP is scarce, and CRP is unable to stimulate transcription at a significant rate, even though no repressor is bound.

CONCEPT CHECK 18.1

1. How does binding of the *trp* corepressor to the *trp* repressor alter repressor function and transcription? What about the binding of the *lac* inducer to the *lac* repressor?
2. Describe the binding of RNA polymerase, repressors, and activators to the *lac* operon when both lactose and glucose are scarce. What is the effect of these scarcities on transcription of the *lac* operon?
3. **WHAT IF?** ➤ A certain mutation in *E. coli* changes the *lac* operator so that the active repressor cannot bind. How would this affect the cell's production of β-galactosidase?

For suggested answers, see Appendix A.

CONCEPT 18.2

Eukaryotic gene expression is regulated at many stages

All organisms, whether prokaryotes or eukaryotes, must regulate which genes are expressed at any given time. Both unicellular organisms and the cells of multicellular organisms continually turn genes on and off in response to signals from their external and internal environments. Regulation of gene expression is also essential for cell specialization in multicellular organisms, which are made up of different types of cells. To perform its own distinct role, each cell type must maintain a specific program of gene expression in which certain genes are expressed and others are not.

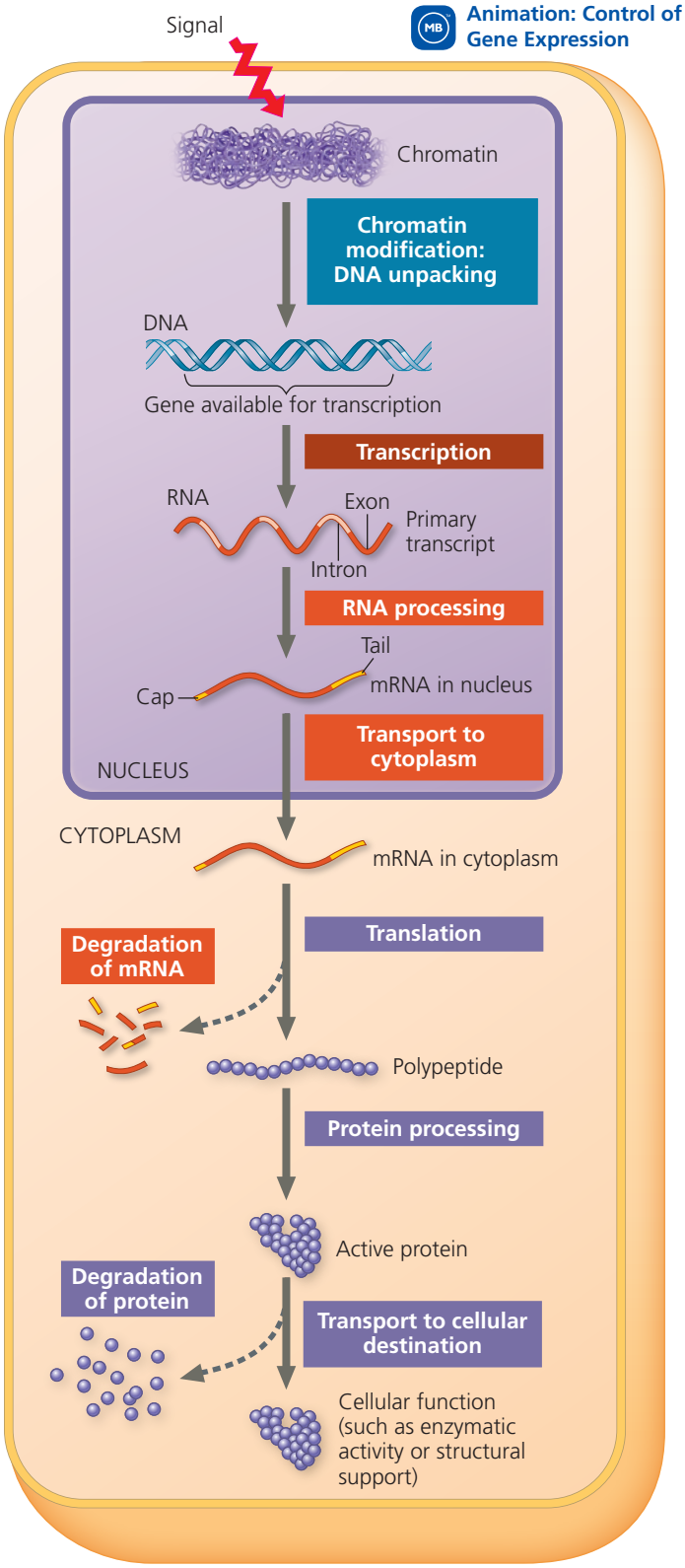
Differential Gene Expression

A typical human cell might express about 20% of its protein-coding genes at any given time. Highly differentiated cells, such as muscle or nerve cells, express an even smaller fraction of their genes. Almost all the cells in a multicellular organism contain an identical genome. (Cells of the immune system are one exception, as you will see in Chapter 47.) A subset of genes is expressed in each cell type; some of these are "housekeeping" genes, expressed by many cell types, while others are unique to that cell type. The uniquely expressed genes allow these cells to carry out their specific function. The differences between cell types, therefore, are due not to different genes being present, but to **differential gene expression**, the expression of different genes by cells with the same genome.

The function of any cell, whether a single-celled eukaryote or a particular cell type in a multicellular organism, depends on the appropriate set of genes being expressed. The transcription factors of a cell must locate the right genes at the right time, a task on a par with finding a needle in a haystack. When gene expression proceeds abnormally, serious imbalances and diseases, including cancer, can arise.

Figure 18.6 summarizes the process of gene expression in a eukaryotic cell, highlighting key stages in the expression of a protein-coding gene. Each stage depicted in Figure 18.6

Figure 18.6 Stages in gene expression that can be regulated in eukaryotic cells. In this diagram, the colored boxes indicate the processes most often regulated; each color indicates the type of molecule that is affected (blue = DNA, red/orange = RNA, purple = protein). The nuclear envelope separating transcription from translation in eukaryotic cells offers an opportunity for post-transcriptional control in the form of RNA processing that is absent in prokaryotes. In addition, eukaryotes have a greater variety of control mechanisms operating before transcription and after translation. A miniature version of this figure accompanies several figures later in the chapter as an orientation diagram.



is a potential control point at which gene expression can be turned on or off, accelerated, or slowed down.

Fifty or so years ago, an understanding of the mechanisms that control gene expression in eukaryotes seemed almost hopelessly out of reach. Since then, new research methods, notably advances in DNA technology (see Chapter 19), have enabled molecular biologists to uncover many details of eukaryotic gene regulation. In all organisms, gene expression is commonly controlled at transcription; regulation at this stage often occurs in response to signals coming from outside the cell, such as hormones or other signaling molecules. For this reason, the term *gene expression* is often equated with transcription for both bacteria and eukaryotes. While this may often be the case for bacteria, the greater complexity of eukaryotic cell structure and function provides opportunities for regulating gene expression at many additional stages (see Figure 18.6). In the remainder of this section, we'll examine some of the important control points of eukaryotic gene expression more closely.

Regulation of Chromatin Structure

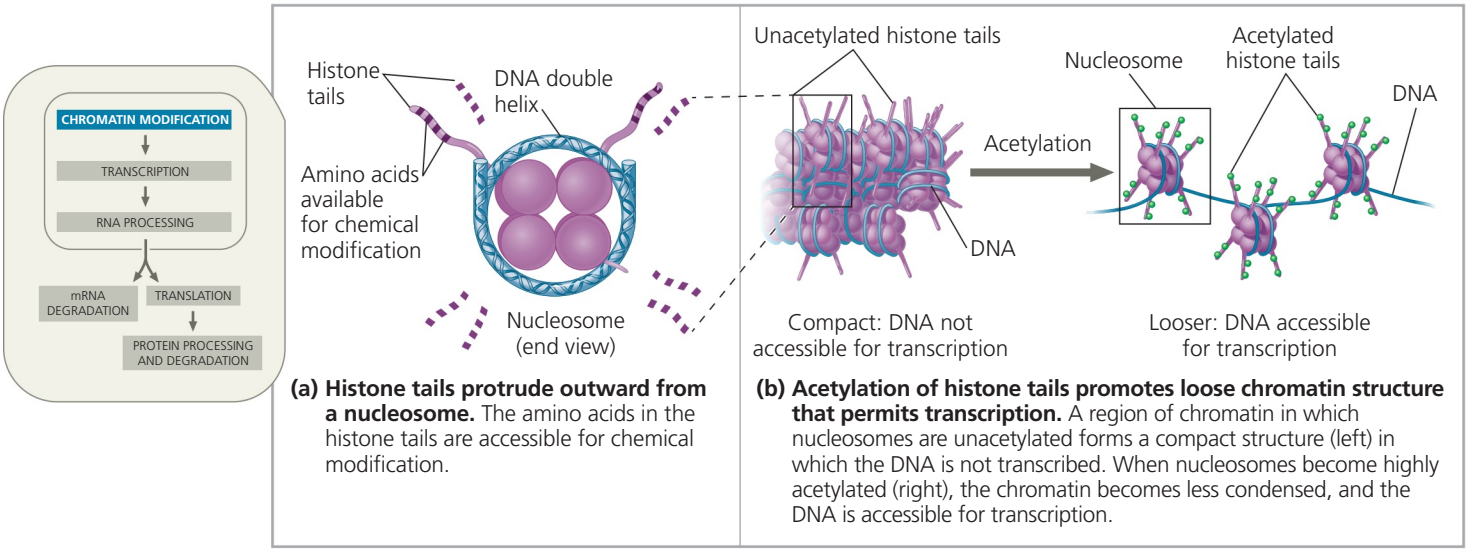
Recall that the DNA of eukaryotic cells is packaged with proteins in an elaborate complex known as chromatin, the basic unit of which is the nucleosome (see Figure 16.22). The structural organization of chromatin not only packs a cell's DNA into a compact form that fits inside the nucleus, but also helps regulate gene expression in several ways. The location of a gene's promoter, relative to both placement of nucleosomes and the sites where the DNA attaches to the chromosome scaffold, can affect whether the gene is transcribed. In addition, genes within heterochromatin, which is highly condensed, are usually not expressed. Lastly, certain chemical modifications to chromatin—both to the histone proteins of the nucleosomes around which DNA is wrapped and to the nucleotides that make up that DNA—can influence chromatin structure and gene expression. Here we examine the effects of these modifications, which are catalyzed by specific enzymes.

Histone Modifications and DNA Methylation

There is abundant evidence that chemical modifications to histones, found in all eukaryotic organisms, play a direct role in the regulation of gene transcription. The N-terminus of each histone protein in a nucleosome protrudes outward from the nucleosome **(Figure 18.7a)**. These so-called *histone tails* are accessible to various modifying enzymes that catalyze the addition or removal of specific chemical groups, such as acetyl (—$COCH_3$), methyl, and phosphate groups (see Figure 4.9). Generally, **histone acetylation**—the addition of an acetyl group to an amino acid in a histone tail—appears to promote transcription by opening up the chromatin structure **(Figure 18.7b)**, while the addition of methyl groups to histones can lead to the condensation of chromatin and reduced transcription. Often, the addition of a particular chemical group may create a new binding site for enzymes that further modify chromatin structure in various ways.

Rather than modifying histone proteins, a different set of enzymes can methylate the DNA itself on certain bases, usually cytosine. Such **DNA methylation** occurs in most plants, animals, and fungi. Long stretches of inactive DNA, such as that of inactivated mammalian X chromosomes (see Figure 15.8), are generally more methylated than regions of actively transcribed DNA (although there are exceptions). On a smaller scale, the DNA of individual genes is usually more heavily methylated in cells in which those genes are not expressed. Removal of the extra methyl groups can turn on some of these genes.

Figure 18.7 A simple model of histone tails and the effect of histone acetylation.
In addition to acetylation, histones can undergo several other types of modifications, such as methylation or phosphorylation. These also help determine the chromatin configuration in a region, sometimes by establishing binding sites for chromatin-modifying enzymes.



(a) Histone tails protrude outward from a nucleosome. The amino acids in the histone tails are accessible for chemical modification.

(b) Acetylation of histone tails promotes loose chromatin structure that permits transcription. A region of chromatin in which nucleosomes are unacetylated forms a compact structure (left) in which the DNA is not transcribed. When nucleosomes become highly acetylated (right), the chromatin becomes less condensed, and the DNA is accessible for transcription.

Once methylated, genes usually stay that way through successive cell divisions in a given individual. At DNA sites where one strand is already methylated, enzymes methylate the correct daughter strand after each round of DNA replication. Methylation patterns are thus passed on to daughter cells, and cells forming specialized tissues keep a chemical record of what occurred during embryonic development. A methylation pattern maintained in this way also accounts for *genomic imprinting* in mammals, where methylation permanently regulates expression of either the maternal or paternal allele of particular genes at the start of development (see Figure 15.17).

Epigenetic Inheritance

The chromatin modifications that we just discussed do not change the DNA sequence, yet they still may be passed along to future generations of cells. Inheritance of traits transmitted by mechanisms not involving the nucleotide sequence itself is called **epigenetic inheritance**. Whereas mutations in the DNA are permanent changes, modifications to the chromatin can be reversed. For example, DNA methylation patterns are largely erased and reestablished during gamete formation.

Researchers are amassing more and more evidence for the importance of epigenetic information in the regulation of gene expression. Epigenetic variations might help explain why one identical twin acquires a genetically based disease, such as schizophrenia, but the other does not, despite their identical genomes. Alterations in normal patterns of DNA methylation are seen in some cancers, where they are associated with inappropriate gene expression. Evidently, enzymes that modify chromatin structure are integral parts of the eukaryotic cell's machinery for regulating transcription.

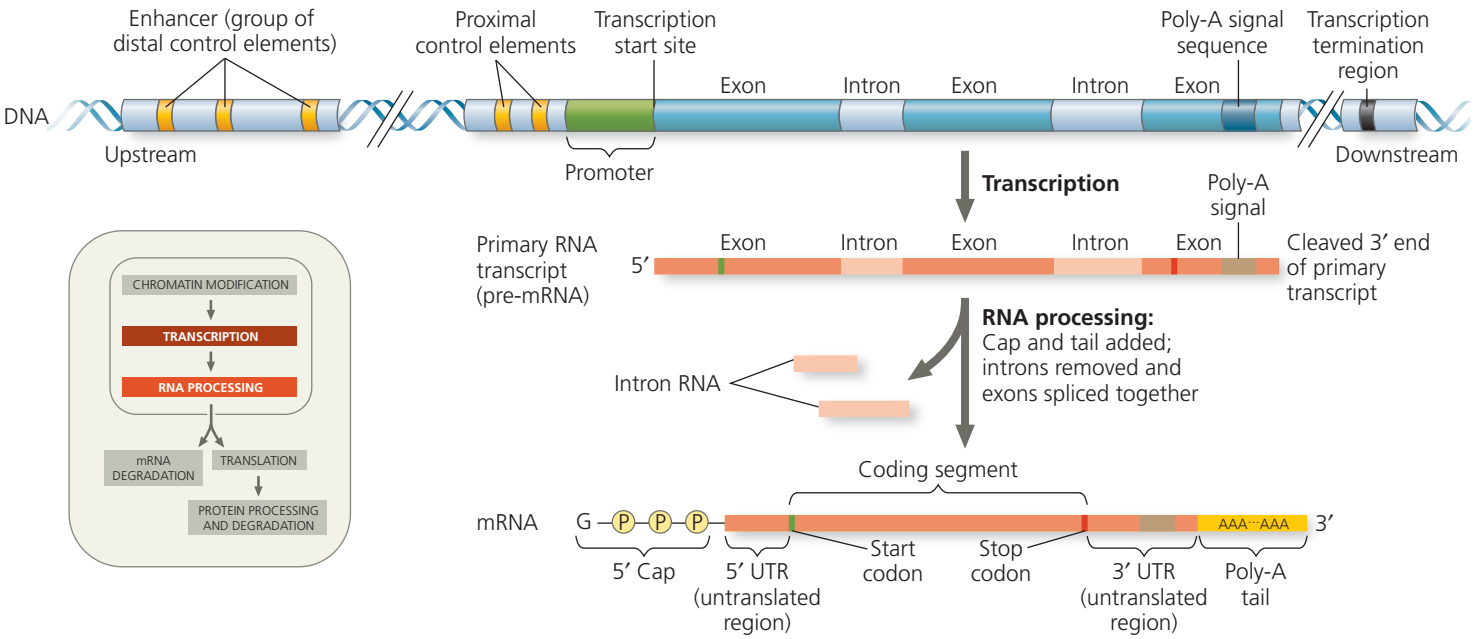
Regulation of Transcription Initiation

Chromatin-modifying enzymes provide initial control of gene expression by making a region of DNA either more or less able to bind the transcription machinery. Once the chromatin of a gene is optimally modified for expression, the initiation of transcription is the next major step at which gene expression is regulated. As in bacteria, the regulation of transcription initiation in eukaryotes involves proteins that bind to DNA and either facilitate or inhibit binding of RNA polymerase. The process is more complicated in eukaryotes, however. Before looking at how eukaryotic cells control their transcription, let's review the structure of a eukaryotic gene.

Organization of a Typical Eukaryotic Gene and Its Transcript

A eukaryotic gene and the DNA elements (segments) that control it are typically organized as shown in **Figure 18.8**, which extends what you learned about eukaryotic genes in Chapter 17. Recall that a cluster of proteins called a *transcription initiation complex* assembles on the promoter sequence

Figure 18.8 A eukaryotic gene and its transcript. Each eukaryotic gene has a promoter—a DNA sequence where RNA polymerase binds and starts transcription, proceeding "downstream." A number of control elements (gold) are involved in regulating the initiation of transcription; these are DNA sequences located near (proximal to) or far from (distal to) the promoter. Distal control elements can be grouped together as enhancers, one of which is shown for this gene. At the other end of the gene, a polyadenylation (poly-A) signal sequence in the last exon of the gene is transcribed into an RNA sequence that signals where the transcript is cleaved and the poly-A tail added. Transcription may continue for hundreds of nucleotides beyond the poly-A signal before terminating. RNA processing of the primary transcript into a functional mRNA involves three steps: addition of the 5′ cap, addition of the poly-A tail, and splicing. In the cell, the 5′ cap is added soon after transcription is initiated, and splicing occurs while transcription is still under way (see Figure 17.11).



at the "upstream" end of the gene (see Figure 17.9). One of these proteins, RNA polymerase II, then proceeds to transcribe the gene, synthesizing a primary RNA transcript (pre-mRNA). RNA processing includes enzymatic addition of a 5′ cap and a poly-A tail, as well as splicing out of introns, to yield a mature mRNA. Associated with most eukaryotic genes are multiple **control elements**, segments of noncoding DNA that serve as binding sites for the proteins called transcription factors, which bind to the control elements and regulate transcription. Control elements on the DNA and the transcription factors that bind to them are critical to the precise regulation of gene expression seen in different cell types.

The Roles of General and Specific Transcription Factors

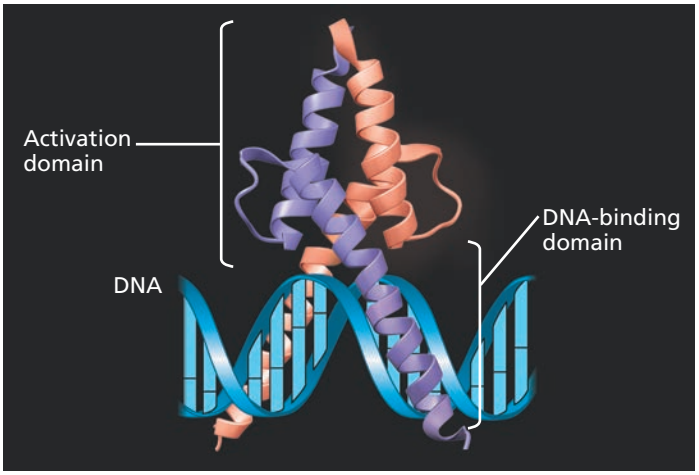
There are two types of transcription factors: General transcription factors act at the promoter of all genes, while some genes require specific transcription factors that bind to control elements that may be close to or farther away from the promoter.

General Transcription Factors at the Promoter To initiate transcription, eukaryotic RNA polymerase requires the assistance of transcription factors. Some transcription factors (such as those illustrated in Figure 17.9) are essential for the transcription of *all* protein-coding genes; therefore, they are often called *general transcription factors*. A few general transcription factors bind to a DNA sequence such as the TATA box within the promoter, but many bind to proteins, including other transcription factors and RNA polymerase II. Protein-protein interactions are crucial to the initiation of eukaryotic transcription. Only when the complete initiation complex has assembled can the polymerase begin to move along the DNA template strand, producing a complementary strand of RNA.

The interaction of general transcription factors and RNA polymerase II with a promoter usually leads to a low rate of initiation and production of few RNA transcripts from genes that are not expressed all the time, but instead are regulated. In eukaryotes, high levels of transcription of these particular genes at the appropriate time and place depend on the interaction of control elements with another set of proteins, which can be thought of as *specific transcription factors*.

Enhancers and Specific Transcription Factors As you can see in Figure 18.8, some control elements, named *proximal control elements*, are located close to the promoter. (Although some biologists consider proximal control elements part of the promoter, in this text we do not.) The more distant *distal control elements*, groupings of which are called **enhancers**, may be thousands of nucleotides upstream or downstream of a gene or even within an intron. A given gene may have multiple enhancers, each active at a different time or in a different cell type or location in the organism. Each enhancer, however, is generally associated with only that gene and no other.

Figure 18.9 The structure of MyoD, a transcriptional activator. The MyoD protein is made up of two polypeptide subunits (purple and salmon) with extensive regions of α helix. Each subunit has one DNA-binding domain (lower half) and one activation domain (upper half). The latter includes binding sites for the other subunit and for other proteins. MyoD is involved in muscle development in vertebrate embryos (see Concept 18.4).

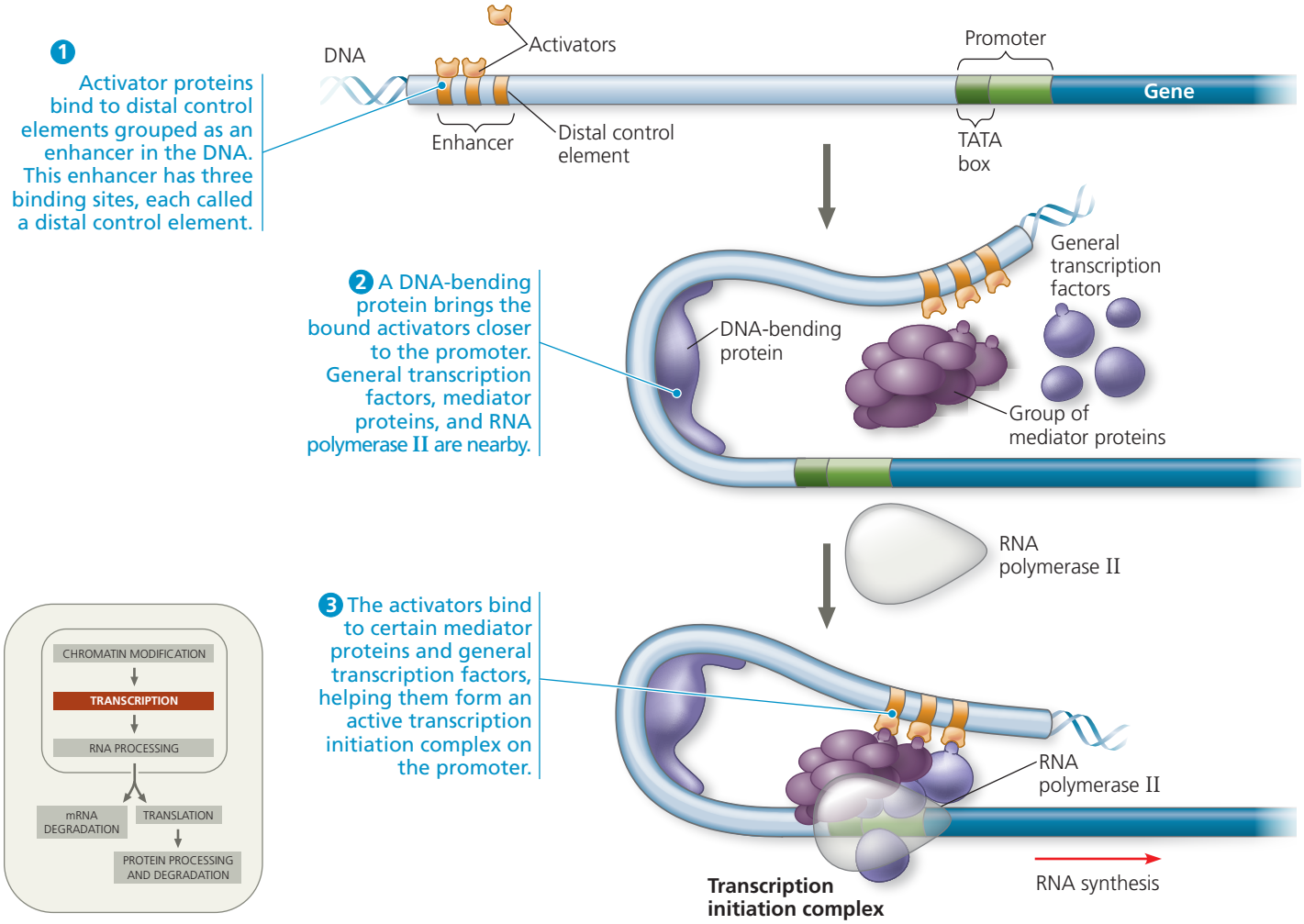


VISUAL SKILLS *Describe how the two functional domains of the MyoD protein relate to the two polypeptide subunits.*

In eukaryotes, the rate of gene expression can be strongly increased or decreased by the binding of specific transcription factors, either activators or repressors, to the control elements of enhancers. Hundreds of transcription activators have been discovered in eukaryotes; the structure of one example is shown in **Figure 18.9**. Researchers have identified two types of structural domains that are commonly found in a large number of activator proteins: a *DNA-binding domain*—a part of the protein's three-dimensional structure that binds to DNA—and one or more *activation domains*. Activation domains bind other regulatory proteins or components of the transcription machinery, facilitating a series of protein-protein interactions that result in enhanced transcription of a given gene.

How can binding of activators to an enhancer located far from the promoter influence transcription? One study shows that the proteins regulating a mouse globin gene contact both the gene's promoter and an enhancer located about 50,000 nucleotides upstream. This and many other studies support the currently accepted model, in which protein-mediated bending of the DNA is thought to bring the bound activators into contact with a group of *mediator proteins*, which in turn interact with general transcription factors at the promoter **(Figure 18.10)**. These protein-protein interactions help assemble and position the initiation complex on the promoter, and allow the promoter and enhancer to come together in a very specific fashion, in spite of what is often a large number of nucleotide pairs between them. In the **Scientific Skills Exercise**, you can work with data from an experiment that identified the control elements in an enhancer of a particular human gene.

Figure 18.10 A model for the action of enhancers and transcription activators. Bending of the DNA by a protein enables enhancers to influence a promoter hundreds or even thousands of nucleotides away. Specific transcription factors called activators bind to the enhancer DNA sequences and then to a group of mediator proteins. These in turn bind to general transcription factors and then RNA polymerase II, thus assembling the transcription initiation complex. These protein-protein interactions lead to correct positioning of the complex on the promoter and the initiation of RNA synthesis. Only one enhancer (with three gold control elements) is shown here, but a gene may have several enhancers that act at different times or in different cell types.



MB **Animation: Regulation of Transcription by Transcriptional Activators and Enhancers**

Specific transcription factors that function as repressors can inhibit gene expression in several different ways. Some repressors bind directly to control element DNA (in enhancers or elsewhere), blocking activator binding. Other repressors interfere with the activator itself so it can't bind the DNA.

In addition to influencing transcription directly, some activators and repressors act indirectly by affecting chromatin structure. Studies using yeast and mammalian cells show that some activators recruit proteins that acetylate histones near the promoters of specific genes, thus promoting transcription (see Figure 18.7). Similarly, some repressors recruit proteins that remove acetyl groups from histones, leading to reduced transcription, a phenomenon referred to as *silencing*. Indeed, recruitment of chromatin-modifying proteins seems to be the most common mechanism of repression in eukaryotic cells.

Combinatorial Control of Gene Activation In eukaryotes, the precise control of transcription depends largely on the binding of activators to DNA control elements. Considering the great number of genes that must be regulated in a typical animal or plant cell, the number of completely different nucleotide sequences found in control elements is surprisingly small. A dozen or so short nucleotide sequences appear again and again in the control elements for different genes. On average, each enhancer is composed of about ten control elements, each of which can

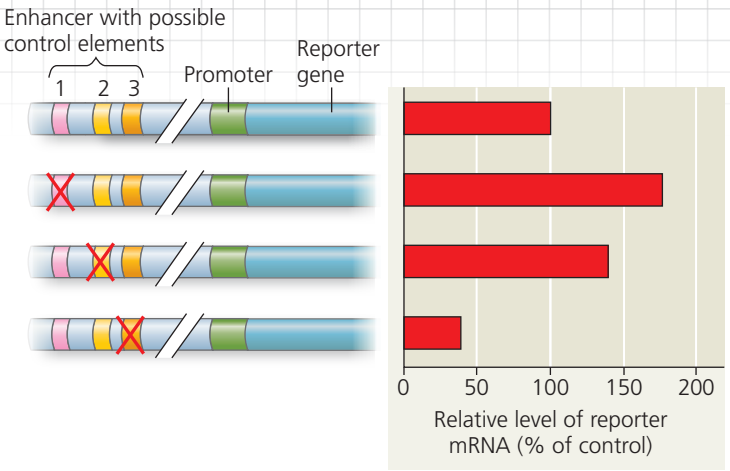
SCIENTIFIC SKILLS EXERCISE

Analyzing DNA Deletion Experiments

What Control Elements Regulate Expression of the *mPGES-1* Gene? The promoter of a gene includes the DNA immediately upstream of the transcription start site, but the control elements that regulate the level of transcription of the gene (grouped in an enhancer) may be thousands of base pairs upstream of the promoter. Because the distance and spacing of control elements make them difficult to identify, scientists begin by deleting possible control elements and measuring the effect on gene expression. In this exercise, you will analyze data obtained from DNA deletion experiments that tested possible control elements for the human gene *mPGES-1*. This gene codes for an enzyme that synthesizes a type of prostaglandin, a chemical made during tissue inflammation.

How the Experiment Was Done The researchers hypothesized that there were three possible control elements in an enhancer region located 8–9 kilobases upstream of the *mPGES-1* gene. Control elements regulate whatever gene is in the appropriate downstream location. Thus, to test the activity of the possible elements, researchers first synthesized molecules of DNA ("constructs") that had the intact enhancer region upstream of a "reporter gene," a gene whose mRNA product could be easily measured experimentally. Next, they made three more DNA constructs, with one of the three proposed control elements deleted in each (see the left side of the figure). The researchers then introduced each DNA construct into a separate human cell culture, where the cells took up the DNA constructs. After 48 hours, the amount of reporter gene mRNA made by the cells was measured. Comparing these amounts allowed researchers to determine if any of the deletions had an effect on expression of the reporter gene, mimicking the effect of deletions on *mPGES-1* gene expression. (The *mPGES-1* gene itself couldn't be used to measure expression levels because the cells express their own *mPGES-1* gene, and the mRNA from that gene would confuse the results.)

Data from the Experiment The diagrams on the left side of the figure show the intact DNA sequence (top) and the three experimental DNA constructs. A red X is located on the possible control element (1, 2, or 3) that was deleted in each experimental DNA construct. The area between the slashes represents the approximately 8 kilobases of DNA located between the promoter and the enhancer region. The horizontal bar graph on the right shows the amount of reporter gene mRNA that was present in each cell culture after 48 hours relative to the amount that was in the culture containing the intact enhancer region (top bar = 100%).



Data from J. N. Walters et al., Regulation of human microsomal prostaglandin E synthase-1 by IL-1b requires a distal enhancer element with a unique role for C/EBPb, *Biochemical Journal* 443:561–571 (2012).

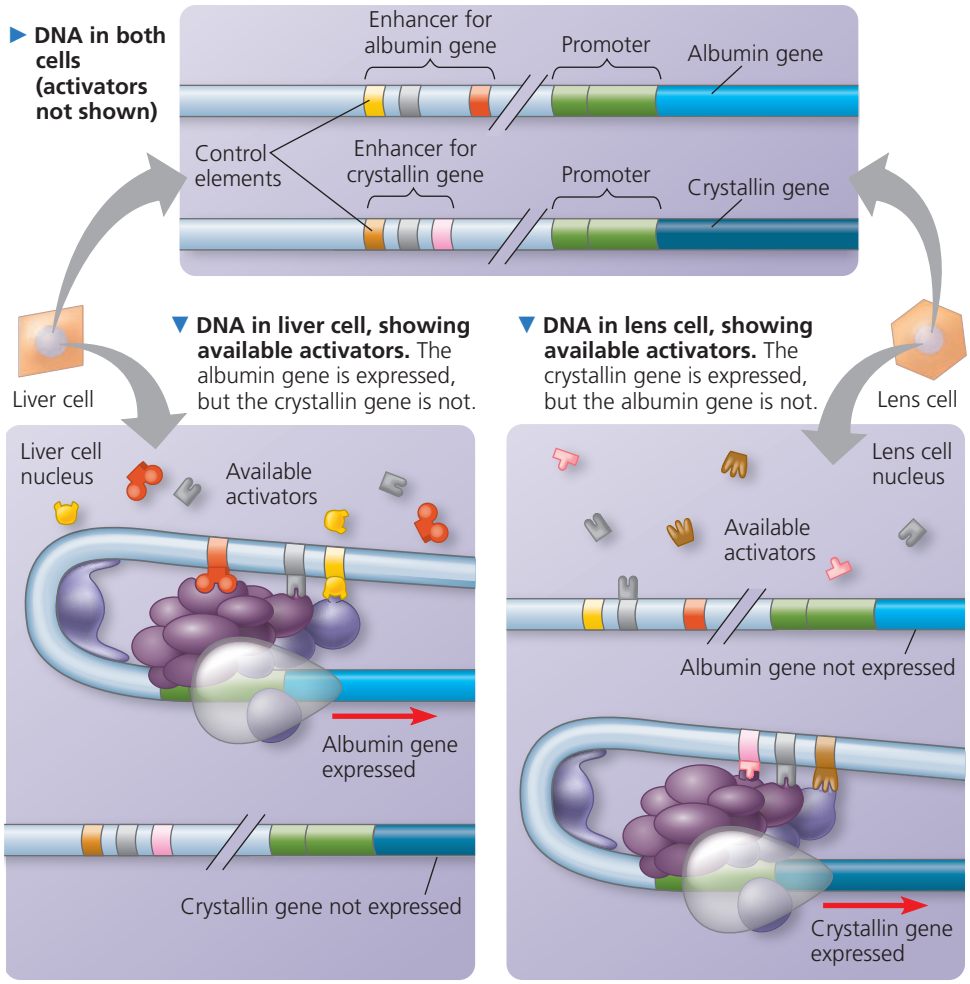
INTERPRET THE DATA

1. (a) What is the independent variable in the graph? (b) What is the dependent variable? (c) What was the control treatment in this experiment? Label it on the diagram.
2. Do the data suggest that any of these possible control elements are actual control elements? Explain.
3. (a) Did deletion of any of the possible control elements cause a *reduction* in reporter gene expression? If so, which one(s), and how can you tell? (b) If loss of a control element causes a reduction in gene expression, what must be the normal role of that control element? Provide a biological explanation for how the loss of such a control element could lead to a reduction in gene expression.
4. (a) Did deletion of any of the possible control elements cause an *increase* in reporter gene expression relative to the control? If so, which one(s), and how can you tell? (b) If loss of a control element causes an increase in gene expression, what must be the normal role of that control element? Propose a biological explanation for how the loss of such a control element could lead to an increase in gene expression.

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

bind only one or two specific transcription factors. It is the particular *combination* of control elements in an enhancer associated with a gene, rather than the presence of a single unique control element, that is important in regulating transcription of the gene.

Even with only a dozen control element sequences available, a very large number of combinations is possible. Each combination of control elements is able to activate transcription only when the appropriate activator proteins are present, which may occur at a precise time during development or in a particular cell type. **Figure 18.11** illustrates how the use of different combinations of just a few control elements can allow differential regulation of transcription in two representative cell types—liver cells and lens cells. This can occur because each cell type contains a different group of activator proteins. How cell types come to differ during this process, even though they all arise from one cell (the fertilized egg), will be explored in Concept 18.4.



➤ Figure 18.11 Cell type–specific transcription. Both liver cells and lens cells have the genes for making the proteins albumin and crystallin, but only liver cells make albumin (a blood protein) and only lens cells make crystallin (the main protein of the lens of the eye). The specific transcription factors made in a cell determine which genes are expressed. In this example, the genes for albumin and crystallin are shown at the top, each with an enhancer made up of three different control elements. Although the enhancers for the two genes both have a gray control element, each enhancer has a unique *combination* of elements. All the activator proteins required for high-level expression of the albumin gene are present in liver cells only (left), whereas the activators needed for expression of the crystallin gene are present in lens cells only (right). For simplicity, we consider only the role of specific transcription factors that are activators here, although repressors may also influence transcription in certain cell types.

VISUAL SKILLS ➤ *Describe the enhancer for the albumin gene in each type of cell. How would the nucleotide sequence of this enhancer in the liver cell compare with that in the lens cell?*

Coordinately Controlled Genes in Eukaryotes

How does the eukaryotic cell deal with a group of genes of related function that need to be turned on or off at the same time? Earlier in this chapter, you learned that in bacteria, such *coordinately controlled* genes are often clustered into an operon, which is regulated by a single promoter and transcribed into a single mRNA molecule. Thus, the genes are expressed together, and the encoded proteins are produced concurrently. With a few exceptions, operons that work in this way have *not* been found in eukaryotic cells.

Eukaryotic genes that are co-expressed, such as genes coding for the enzymes of a metabolic pathway, are typically scattered over different chromosomes. Here, coordinate gene expression depends on every gene of a dispersed group having a specific combination of control elements. Activator proteins in the nucleus that recognize the control elements bind to them, promoting simultaneous transcription of the genes, no matter where they are in the genome.

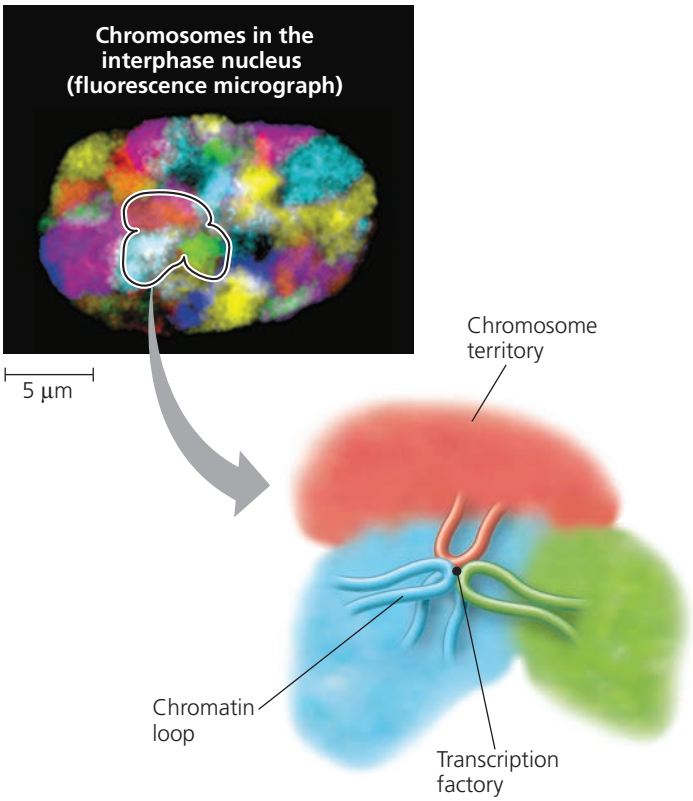
Coordinate control of dispersed genes in a eukaryotic cell often occurs in response to chemical signals from outside the cell. A steroid hormone, for example, enters a cell and binds to a specific intracellular receptor protein, forming a hormone-receptor complex that serves as a transcription activator (see Figure 9.9). Every gene whose transcription is stimulated by a given steroid hormone, regardless of its chromosomal location, has a control element recognized by that hormone-receptor complex. This is how estrogen activates a group of genes that stimulate cell division in uterine cells, preparing the uterus for pregnancy.

Many signaling molecules, such as nonsteroid hormones and growth factors, bind to receptors on a cell's surface and never actually enter the cell. Such molecules can control gene expression indirectly by triggering signal transduction pathways that activate particular transcription factors (see Figure 9.15). Coordinate regulation in such pathways is the same as for steroid hormones: Genes with the same sets of control elements are activated by the same chemical signals. Because this system for coordinating gene regulation is so widespread, biologists think that it probably arose early in evolutionary history.

Nuclear Architecture and Gene Expression

You saw in Figure 16.23b that each chromosome in the interphase nucleus occupies a distinct territory. The chromosomes are not completely isolated, however. Recently, *chromosome conformation capture* (*3C*) techniques have been developed that allow researchers to cross-link and identify regions of

Figure 18.12 Chromosomal interactions in the interphase nucleus. Although each chromosome has its own territory (see Figure 16.23b), loops of chromatin may extend into other sites in the nucleus. Some of these sites are transcription factories that are occupied by multiple chromatin loops from the same chromosome (blue loops) or other chromosomes (red and green loops).



chromosomes that associate with each other during interphase. These studies reveal that loops of chromatin extend from individual chromosomal territories into specific sites in the nucleus **(Figure 18.12)**. Different loops from the same chromosome and loops from other chromosomes may congregate in such sites, some of which are rich in RNA polymerases and other transcription-associated proteins. Like a recreation center that draws members from many different neighborhoods, these so-called *transcription factories* are thought to be areas specialized for a common function.

The old view that the nuclear contents are like a bowl of amorphous chromosomal spaghetti has given way to a new model of a nucleus with a defined architecture and regulated movements of chromatin. Several lines of evidence suggest that genes that are not being expressed are located in the outer edges of the nucleus, while those that are being expressed are found in its interior region. Relocation of particular genes from their chromosomal territories to transcription factories in the interior may be part of the process of readying genes for transcription. How long an individual transcription factory may last has not yet been established. In 2014, the National Institutes of Health announced funding for a new "4D Nucleome" program, which aims to investigate the many fascinating questions addressed by this exciting area of current research.

Mechanisms of Post-Transcriptional Regulation

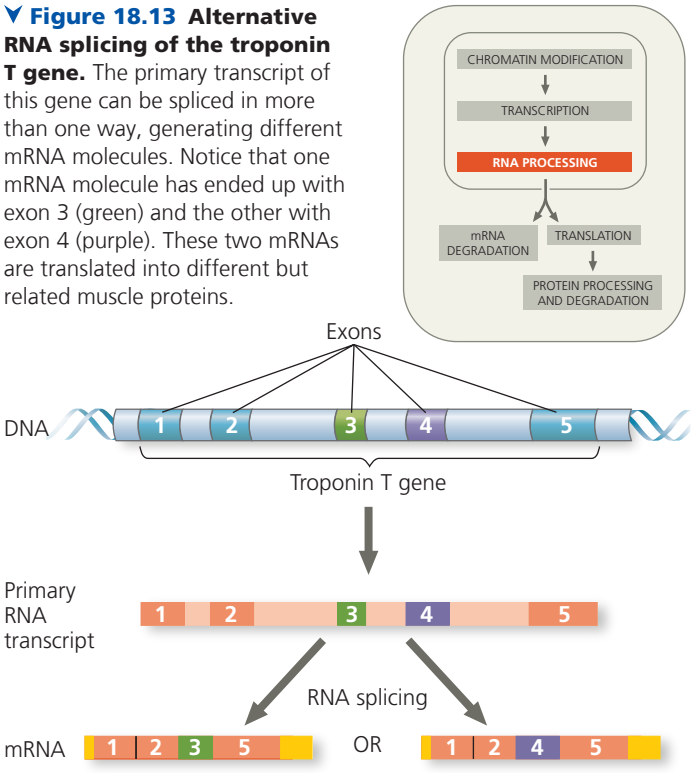
Transcription alone does not constitute gene expression. The expression of a protein-coding gene is ultimately measured by the amount of functional protein a cell makes, and much happens between the synthesis of the RNA transcript and the activity of the protein in the cell. Many regulatory mechanisms operate at the various stages after transcription (see Figure 18.6). These mechanisms allow a cell to rapidly fine-tune gene expression in response to environmental changes without altering its transcription patterns. Here we discuss how cells can regulate gene expression once a gene has been transcribed.

RNA Processing

RNA processing in the nucleus and the export of mature RNA to the cytoplasm provide opportunities for regulating gene expression not available in prokaryotes. One example of regulation at the RNA-processing level is **alternative RNA splicing**, in which different mRNA molecules are produced from the same primary transcript, depending on which RNA segments are treated as exons and which as introns. Regulatory proteins specific to a cell type control intron-exon choices by binding to regulatory sequences within the primary transcript.

A simple example of alternative RNA splicing is shown in **Figure 18.13** for the troponin T gene, which encodes two different (though related) proteins. Other genes code for many more possible products. For instance, researchers have found a *Drosophila* gene with enough alternatively spliced exons to

Figure 18.13 Alternative RNA splicing of the troponin T gene. The primary transcript of this gene can be spliced in more than one way, generating different mRNA molecules. Notice that one mRNA molecule has ended up with exon 3 (green) and the other with exon 4 (purple). These two mRNAs are translated into different but related muscle proteins.



generate about 19,000 membrane proteins that have different extracellular domains. At least 17,500 (94%) of the alternative mRNAs are actually synthesized. Each developing nerve cell in the fly appears to synthesize a different form of the protein, which acts as unique identifier on the cell surface and helps prevent excessive overlap of nerve cells during development of the nervous system.

It is clear that alternative RNA splicing can significantly expand the repertoire of a eukaryotic genome. In fact, alternative splicing was proposed as one explanation for the surprisingly low number of human genes counted when the human genome was sequenced. The number of human genes was found to be similar to that of a soil worm (nematode), a mustard plant, or a sea anemone. This discovery prompted questions about what, if not the number of genes, accounts for the more complex morphology (external form) of humans. It turns out that more than 90% of human protein-coding genes probably undergo alternative splicing. Thus, the extent of alternative splicing greatly multiplies the number of possible human proteins, which may be better correlated with complexity of form.

Initiation of Translation and mRNA Degradation

Translation is another opportunity for regulating gene expression; it occurs most commonly at the initiation stage (see Figure 17.19). For some mRNAs, the initiation of translation can be blocked by regulatory proteins that bind to specific sequences or structures within the untranslated region (UTR) at the 5′ or 3′ end, preventing the attachment of ribosomes. (Recall from Concept 17.3 that both the 5′ cap and the poly-A tail of an mRNA molecule are important for ribosome binding.)

Alternatively, translation of *all* the mRNAs in a cell may be regulated simultaneously. In a eukaryotic cell, such "global" control usually involves the activation or inactivation of one or more of the protein factors required to initiate translation. This mechanism plays a role in starting translation of mRNAs that are stored in eggs. Just after fertilization, translation is triggered by the sudden activation of translation initiation factors. The response is a burst of synthesis of the proteins encoded by the stored mRNAs. Some plants and algae store mRNAs during periods of darkness; light then triggers the reactivation of the translational apparatus.

The life span of mRNA molecules in the cytoplasm is important in determining the pattern of protein synthesis in a cell. Bacterial mRNA molecules typically are degraded by enzymes within a few minutes of their synthesis. This short life span of mRNAs is one reason bacteria can change their patterns of protein synthesis so quickly in response to environmental changes. In contrast, mRNAs in multicellular eukaryotes typically survive for hours, days, or even weeks. For instance, the mRNAs for the hemoglobin polypeptides (α-globin and β-globin) in developing red blood cells are unusually stable, and these long-lived mRNAs are translated repeatedly in red blood cells.

Nucleotide sequences that affect how long an mRNA remains intact are often found in the untranslated region at the 3′ end of the molecule (see Figure 18.8). In one experiment, researchers transferred such a sequence from the short-lived mRNA for a growth factor to the 3′ end of a normally stable globin mRNA. The globin mRNA was quickly degraded.

During the past few years, other mechanisms that degrade or block expression of mRNA molecules have come to light. They involve a group of newly discovered RNA molecules that regulate gene expression at several levels, as we'll discuss shortly.

Protein Processing and Degradation

The final opportunities for controlling gene expression occur after translation. Often, eukaryotic polypeptides must be processed to yield functional protein molecules. For instance, cleavage of the initial insulin polypeptide (pro-insulin) forms the active hormone. In addition, many proteins undergo chemical modifications that make them functional. Regulatory proteins are commonly activated or inactivated by the reversible addition of phosphate groups (see Figure 9.10), and proteins destined for the surface of animal cells acquire sugars (see Figure 7.12). Cell-surface proteins and many others must also be transported to target destinations in the cell in order to function (see Figure 17.22). Regulation might occur at any of the steps involved in modifying or transporting a protein.

Finally, the length of time each protein functions in the cell is strictly regulated by selective degradation. Many proteins, such as the cyclins involved in regulating the cell cycle, must be relatively short-lived if the cell is to function appropriately (see Figure 12.16). To mark a protein for destruction, the cell commonly attaches molecules of a small protein called ubiquitin to the protein. Giant protein complexes called proteasomes then recognize the ubiquitin-tagged proteins and degrade them.

Animation: Post-Transcriptional Control Mechanisms

CONCEPT CHECK 18.2

1. In general, what are the effects of histone acetylation and DNA methylation on gene expression?
2. **MAKE CONNECTIONS ➤** Speculate about whether the same enzyme could methylate both a histone and a DNA base. (See Concept 5.4.)
3. Compare the roles of general and specific transcription factors in regulating gene expression.
4. Once mRNA encoding a particular protein reaches the cytoplasm, what are four mechanisms that can regulate the amount of the protein that is active in the cell?
5. **WHAT IF? ➤** Suppose you compared the nucleotide sequences of the distal control elements in the enhancers of three genes that are expressed only in muscle cells. What would you expect to find? Why?

For suggested answers, see Appendix A.

CONCEPT 18.3

Noncoding RNAs play multiple roles in controlling gene expression

Genome sequencing has revealed that protein-coding DNA accounts for only 1.5% of the human genome and a similarly small percentage of the genomes of many other multicellular eukaryotes. A very small fraction of the non-protein-coding DNA consists of genes for RNAs such as ribosomal RNA and transfer RNA. Until recently, scientists assumed that most of the remaining DNA was not transcribed, thinking that since it didn't specify proteins or the few known types of RNA, such DNA didn't contain meaningful genetic information—in fact, it was called "junk DNA." However, a flood of recent data has contradicted this idea. For example, a massive study of the entire human genome showed that roughly 75% of the genome is transcribed at some point in any given cell. Introns account for only a fraction of this transcribed, nontranslated RNA. These and other results suggest that a significant amount of the genome may be transcribed into non-protein-coding RNAs—also called *noncoding RNAs*, or *ncRNAs*—including a variety of small RNAs. Researchers are uncovering more evidence of the biological roles of these ncRNAs every day.

Biologists are excited about these discoveries, which have revealed a large and diverse population of RNA molecules in the cell that play crucial roles in regulating gene expression—but have gone largely unnoticed until fairly recently. Our long-standing view that because mRNAs code for proteins, they are the most important RNAs functioning in the cell demands revision. This represents a major shift in the thinking of biologists, one that you are witnessing as students entering this field of study.

Effects on mRNAs by MicroRNAs and Small Interfering RNAs

Regulation by both small and large ncRNAs occurs at several points in the pathway of gene expression, including mRNA translation and chromatin modification. We'll examine two types of small ncRNAs, the importance of which was acknowledged when their discovery was the focus of the 2006 Nobel Prize in Physiology or Medicine, which was awarded for work completed only eight years earlier.

Since 1993, a number of research studies have uncovered **microRNAs (miRNAs)**—small, single-stranded RNA molecules capable of binding to complementary sequences in mRNA molecules. A longer RNA precursor is processed by cellular enzymes into an miRNA, a single-stranded RNA of about 22 nucleotides that forms a complex with one or more proteins **(Figure 18.14)**. The miRNA allows the complex to bind to any mRNA molecule with at least 7 or 8 nucleotides of complementary sequence. The miRNA-protein complex then degrades the target mRNA or, less often, simply blocks its translation. There are approximately 1,500 genes for miRNAs in the human genome, and biologists estimate that expression of at least one-half of all human genes may be regulated by miRNAs, a remarkable figure given that the existence of miRNAs was unknown 25 years ago.

Another class of small RNAs, similar in size and function to miRNAs, is called **small interfering RNAs (siRNAs)**. Both miRNAs and siRNAs can associate with the same proteins, producing similar results. In fact, if siRNA precursor RNA molecules are injected into a cell, the cell's machinery can process them into siRNAs that turn off expression of genes with related sequences, similarly to how miRNAs function. The distinction between miRNAs and siRNAs is based on subtle differences in the structure of their precursors, which in both cases are RNA molecules that are mostly double-stranded. The blocking of gene expression by siRNAs is referred to as **RNA interference (RNAi)**, and it is used in the laboratory as a means of disabling specific genes to investigate their function.

How did the RNAi pathway evolve? As you will learn in Concept 26.2, some viruses have double-stranded RNA genomes. Given that the cellular RNAi pathway can process double-stranded RNAs into homing devices that lead to destruction of related RNAs, some scientists think that this pathway may have evolved as a natural defense against

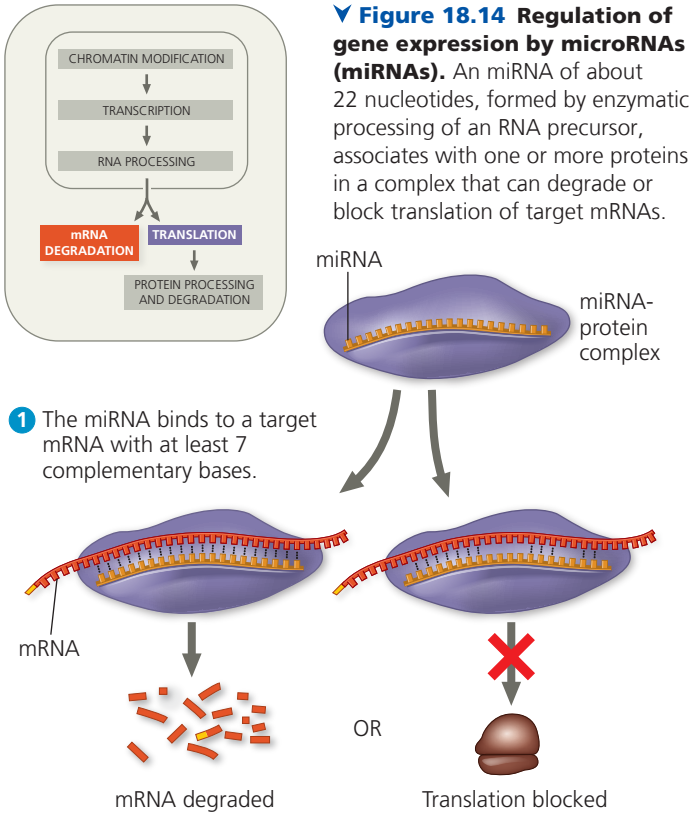


Figure 18.14 Regulation of gene expression by microRNAs (miRNAs). An miRNA of about 22 nucleotides, formed by enzymatic processing of an RNA precursor, associates with one or more proteins in a complex that can degrade or block translation of target mRNAs.

infection by such viruses. However, the fact that RNAi can also affect the expression of nonviral cellular genes may reflect a different evolutionary origin for the RNAi pathway. Moreover, many species, including mammals, apparently produce their own long, double-stranded RNA precursors to small RNAs such as siRNAs. Once produced, these RNAs can interfere with gene expression at stages other than translation, as we'll discuss next.

Chromatin Remodeling and Effects on Transcription by ncRNAs

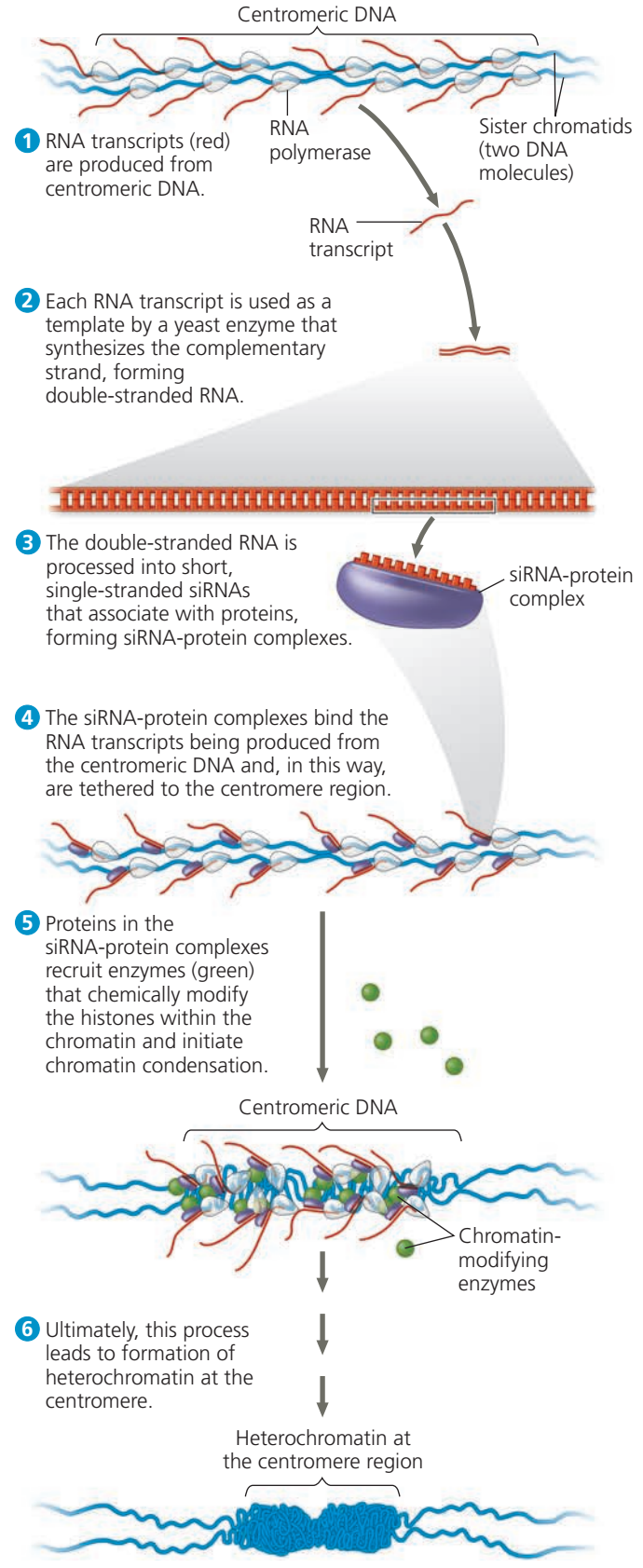
In addition to regulating mRNAs, some ncRNAs act to bring about remodeling of chromatin structure. One example occurs during formation of heterochromatin at the centromere, as studied in a species of yeast.

In the S phase of the cell cycle, the centromeric regions of DNA must be loosened for chromosomal replication and then re-condensed into heterochromatin in preparation for mitosis. In some yeasts, siRNAs produced by the yeast cells themselves are required to re-form the heterochromatin at the centromeres. A model for how this happens is shown in **Figure 18.15**. Exactly how the process starts is still under debate, but biologists agree on the general idea: The siRNA system in yeast interacts with other ncRNAs and with chromatin-modifying enzymes to condense the centromere chromatin into heterochromatin. Under most conditions in mammalian cells, siRNAs are not known to occur, and the mechanism for centromere DNA condensation is not yet understood. However, it may turn out to involve other small ncRNAs.

A recently discovered class of small ncRNAs is called *piwi-interacting RNAs*, or *piRNAs*. These RNAs also induce formation of heterochromatin, blocking expression of some parasitic DNA elements in the genome known as transposons. (Transposons are discussed in Concepts 20.4 and 20.5.) Usually 24–31 nucleotides in length, piRNAs are processed from a longer, single-stranded RNA precursor. They play an indispensable role in the germ cells of many animal species, where they appear to help reestablish appropriate methylation patterns in the genome during gamete formation.

Researchers have also found a relatively large number of **long noncoding RNAs (lncRNAs)**, ranging from 200 to hundreds of thousands of nucleotides in length, that are expressed at significant levels in specific cell types at particular times. One such lncRNA is responsible for X chromosome inactivation, which, in most female mammals, prevents expression of genes located on one of the X chromosomes (see Figure 15.8). In this case, lncRNAs—transcripts of the *XIST* gene located on the

Figure 18.15 Condensation of chromatin at the centromere. In one type of yeast, siRNAs and longer noncoding RNAs cooperate in the pathway that leads to re-formation of highly condensed heterochromatin at the centromere of each chromatid after DNA replication.



chromosome to be inactivated—bind back to and coat that chromosome, and this binding leads to condensation of the entire chromosome into heterochromatin.

The cases described above involve chromatin remodeling in large regions of the chromosome. Because chromatin structure affects transcription and thus gene expression, RNA-based regulation of chromatin structure is sure to play an important role in gene regulation. Additionally, some experimental evidence supports the idea of an alternate role for lncRNAs in which they can act as a scaffold, bringing together DNA, proteins, and other RNAs into complexes. These associations may act either to condense chromatin or, in some cases, to help bring the enhancer of a gene together with mediator proteins and the gene's promoter, activating gene expression in a more direct fashion.

The Evolutionary Significance of Small ncRNAs

EVOLUTION Small ncRNAs can regulate gene expression at multiple steps and in many ways. While this section has focused on ncRNAs in eukaryotes, small ncRNAs are also used by bacteria as a defense system, called the CRISPR-Cas9 system, against viruses that infect them. (You'll learn more about this in Concept 26.2.) The use of ncRNAs thus evolved long ago, but we don't yet know how bacterial ncRNAs are related to those of eukaryotes.

What about the evolutionary significance of small eukaryotic ncRNAs? In general, extra levels of gene regulation might allow evolution of a higher degree of complexity of form. The versatility of miRNA regulation has therefore led some biologists to hypothesize that an increase in the number of different miRNAs specified by the genome of a given species has allowed morphological complexity to increase over evolutionary time. While this hypothesis is still being evaluated, it is logical to expand the discussion to include all small ncRNAs. Exciting new techniques for rapidly sequencing genomes have allowed biologists to begin asking how many genes for ncRNAs are present in the genome of any given species. A survey of different species supports the notion that siRNAs evolved first, followed by miRNAs and later piRNAs, which are found only in animals. And while there are hundreds of types of miRNAs, there appear to be 60,000 or so types of piRNAs, allowing the potential for very sophisticated gene regulation by piRNAs.

Given the extensive functions of ncRNAs, it is not surprising that many of the ncRNAs characterized thus far play important roles in embryonic development—the topic we turn to in the next section. Embryonic development is perhaps the ultimate example of precisely regulated gene expression.

CONCEPT CHECK 18.3

1. Compare miRNAs and siRNAs, including their functions.
2. **WHAT IF?** ➤ Suppose the mRNA being degraded in Figure 18.14 coded for a protein that promotes cell division in a multicellular organism. What would happen if a mutation disabled the gene for the miRNA that triggers this degradation?
3. **MAKE CONNECTIONS** ➤ Inactivation of one of the X chromosomes in female mammals involves lncRNA called *XIST* RNA, mentioned in this section and in Concept 15.2. Describe transcription and binding of *XIST* RNA, then suggest a model for how it initiates Barr body formation.

For suggested answers, see Appendix A.

CONCEPT 18.4

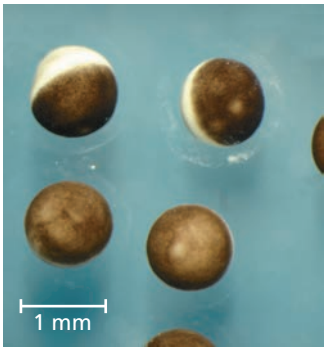
A program of differential gene expression leads to the different cell types in a multicellular organism

In the embryonic development of multicellular organisms, a fertilized egg (a zygote) gives rise to cells of many different types, each with a different structure and corresponding function. Typically, cells are organized into tissues, tissues into organs, organs into organ systems, and organ systems into the whole organism. Thus, any developmental program must produce cells of different types that form higher-level structures arranged in a particular way in three dimensions. The processes that occur during development in plants and animals are detailed in Chapters 35 and 46, respectively. In this chapter, we focus on the program of regulation of gene expression that orchestrates development, using a few animal species as examples.

A Genetic Program for Embryonic Development

The photos in **Figure 18.16** illustrate the dramatic difference between a frog zygote (fertilized egg) and the tadpole it

Figure 18.16 From fertilized egg to animal: What a difference four days makes. It takes just four days for cell division, differentiation, and morphogenesis to transform each of the fertilized frog eggs shown in **(a)** into a tadpole like the one in **(b)**.



(a) Fertilized eggs of a frog

(b) Newly hatched tadpole

becomes. This remarkable transformation results from three interrelated processes: cell division, cell differentiation, and morphogenesis. Through a succession of mitotic cell divisions, the zygote gives rise to a large number of cells. Cell division alone, however, would merely produce a great ball of identical cells, nothing like a tadpole. During embryonic development, cells not only increase in number, but also undergo cell **differentiation**, the process by which cells become specialized in structure and function. Moreover, the different kinds of cells are not randomly distributed but are organized into tissues and organs in a particular three-dimensional arrangement. The physical processes that give an organism its shape constitute **morphogenesis**, the development of the form of an organism and its structures.

All three processes are rooted in cellular behavior. Even morphogenesis, the shaping of the organism, can be traced back to changes in the shape, motility, and other characteristics of the cells that make up various regions of the embryo. As you have seen, the activities of a cell depend on the genes it expresses and the proteins it produces. Almost all cells in an organism have the same genome; therefore, differential gene expression results from the genes being regulated differently in each cell type.

In Figure 18.11, you saw a simplified view of how differential gene expression occurs in two cell types, a liver cell and a lens cell. Each of these fully differentiated cells has a particular mix of specific activators that turn on the collection of genes whose products are required in the cell. The fact that both cells arose through a series of mitoses from a common fertilized egg inevitably leads to a question: How do different sets of activators come to be present in the two cells?

It turns out that materials placed into the egg by maternal cells set up a sequential program of gene regulation that is carried out as embryonic cells divide, and this program coordinates cell differentiation during embryonic development. To understand how this works, we will consider two basic developmental processes. First, we'll explore how cells that arise from early embryonic mitoses develop the differences that start each cell along its own differentiation pathway. Second, we'll see how cellular differentiation leads to one particular cell type, using muscle development as an example.

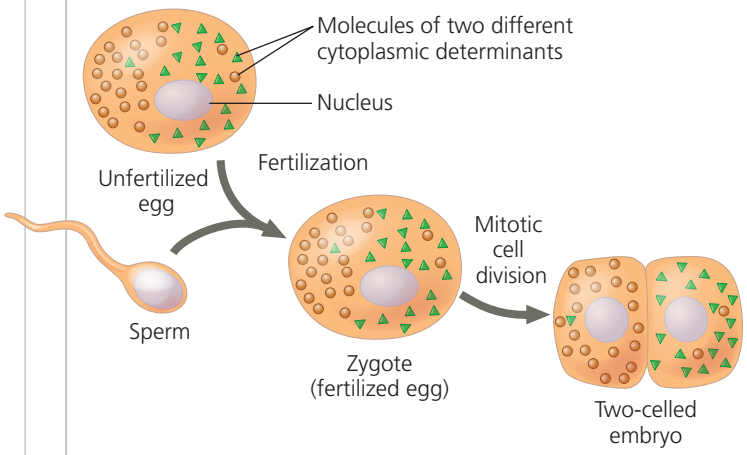
Cytoplasmic Determinants and Inductive Signals

What generates the first differences among cells in an early embryo? And what controls the differentiation of all the various cell types as development proceeds? By this point in the chapter, you can probably deduce the answer: The specific genes expressed in any particular cell of a developing organism determine its path. Two sources of information, used to varying extents in different species, "tell" a cell which genes to express at any given time during embryonic development.

One important source of information early in development is the egg's cytoplasm, which contains both RNA and proteins encoded by the mother's DNA. The cytoplasm of an unfertilized egg is not homogeneous. Messenger RNA, proteins, other substances, and organelles are distributed unevenly in the unfertilized egg, and this unevenness has a profound impact on the development of the future embryo in many species. Maternal substances in the egg that influence the course of early development are called **cytoplasmic determinants** (Figure 18.17a). After fertilization, early mitotic divisions distribute the zygote's

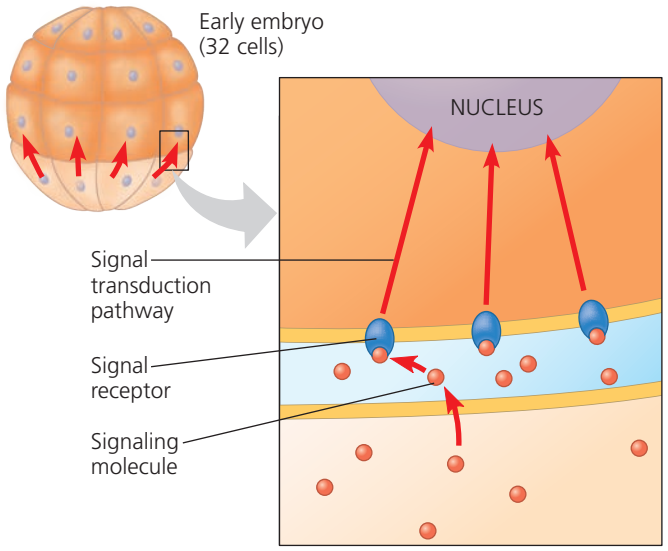
Figure 18.17 Sources of developmental information for the early embryo.

(a) Cytoplasmic determinants in the egg



The unfertilized egg has molecules in its cytoplasm, encoded by the mother's genes, that influence development. Many of these cytoplasmic determinants, like the two shown here, are unevenly distributed in the egg. After fertilization and mitotic division, the cell nuclei of the embryo are exposed to different sets of cytoplasmic determinants and, as a result, express different genes.

(b) Induction by nearby cells



Cells at the bottom of the early embryo are releasing molecules that signal (induce) nearby cells to change their gene expression.

cytoplasm into separate cells. The nuclei of these cells may thus be exposed to different cytoplasmic determinants, depending on which portions of the zygotic cytoplasm a cell received. The combination of cytoplasmic determinants in a cell helps determine its developmental fate by regulating expression of the cell's genes during the course of cell differentiation.

The other major source of developmental information, which becomes increasingly important as the number of embryonic cells increases, is the environment around a particular cell. Most influential are the signals conveyed to an embryonic cell from other embryonic cells in the vicinity, including contact with cell-surface molecules on neighboring cells and the binding of growth factors secreted by neighboring cells (see Concept 9.1). Such signals cause changes in the target cells, a process called **induction** **(Figure 18.17b)**. The molecules passing along these signals within the target cell are cell-surface receptors and other signaling pathway proteins. In general, the signaling molecules send a cell down a specific developmental path by causing changes in its gene expression that eventually result in observable cellular changes. Thus, interactions between embryonic cells help induce differentiation into the many specialized cell types making up a new organism.

Sequential Regulation of Gene Expression During Cellular Differentiation

The earliest changes that set a cell on its path to specialization are subtle ones, showing up only at the molecular level. Before biologists knew much about the molecular changes occurring in embryos, they coined the term **determination** to refer to the point at which an embryonic cell is irreversibly committed to becoming a particular cell type. Once it has undergone determination, an embryonic cell can be experimentally placed in another location in the embryo and it will still differentiate into the cell type that is its normal fate. Differentiation, then, is the process by which a cell attains its determined fate. As the tissues and organs of an embryo develop and their cells differentiate, the cells become more noticeably different in structure and function.

Today we understand determination in terms of molecular changes. The outcome of determination, observable cell differentiation, is marked by the expression of genes for *tissue-specific proteins*. These proteins are found only in a specific cell type and give the cell its characteristic structure and function. The first evidence of differentiation is the appearance of mRNAs for these proteins. Eventually, differentiation is observable with a microscope as changes in cellular structure. On the molecular level, different sets of genes are sequentially expressed in a regulated manner as new cells arise from division of their precursors. A number of the steps in gene expression may be regulated during differentiation, transcription being the most common. In the fully differentiated cell, transcription remains the principal regulatory point for maintaining appropriate gene expression.

Differentiated cells are specialists at making tissue-specific proteins. For example, as a result of transcriptional regulation, liver cells specialize in making albumin, and lens cells specialize in making crystallin (see Figure 18.11). Skeletal muscle cells in vertebrates are another instructive example. Each of these cells is a long fiber containing many nuclei within a single plasma membrane. Skeletal muscle cells have high concentrations of muscle-specific versions of the contractile proteins myosin and actin, as well as membrane receptor proteins that detect signals from nerve cells.

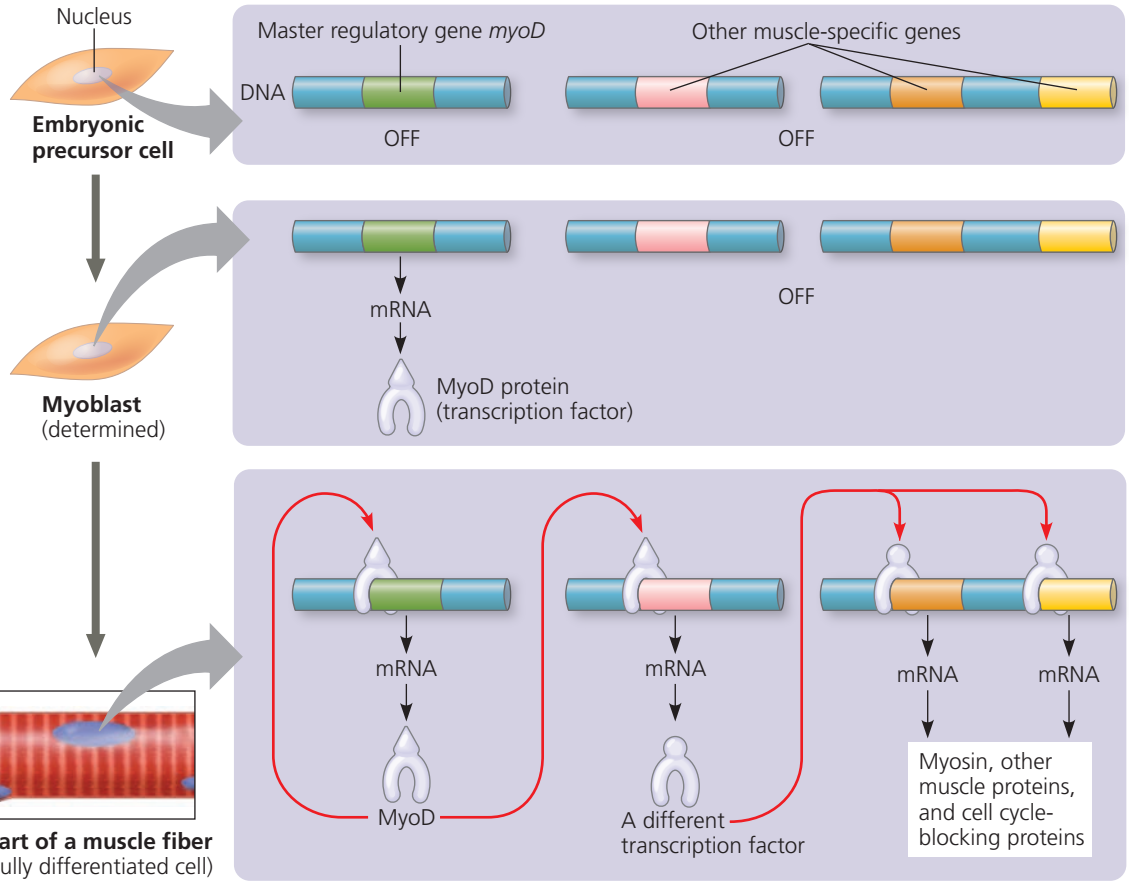
Muscle cells develop from embryonic precursor cells that have the potential to develop into a number of cell types, including cartilage cells and fat cells, but particular conditions commit them to becoming muscle cells. Although the committed cells appear unchanged under the microscope, determination has occurred, and they are now *myoblasts*. Eventually, myoblasts start to churn out large amounts of muscle-specific proteins and fuse to form mature, elongated, multinucleate skeletal muscle cells.

Researchers have worked out what happens at the molecular level during muscle cell determination. To do so, they grew embryonic precursor cells in culture and analyzed them using molecular techniques you will learn about in Concepts 19.1 and 19.2. In a series of experiments, they isolated different genes, caused each to be expressed in a separate embryonic precursor cell, and then looked for differentiation into myoblasts and muscle cells. In this way, they identified several so-called "master regulatory genes" whose protein products commit the cells to becoming skeletal muscle. Thus, in the case of muscle cells, the molecular basis of determination is the expression of one or more of these master regulatory genes.

To understand more about how determination occurs in muscle cell differentiation, let's focus on the master regulatory gene called *myoD*. The *myoD* gene deserves its designation as a master regulatory gene. Researchers have shown that the MyoD protein it encodes is capable of changing some kinds of fully differentiated nonmuscle cells, such as fat cells and liver cells, into muscle cells. Why doesn't MyoD work on *all* kinds of cells? One likely explanation is that activation of muscle-specific genes is not solely dependent on MyoD but requires a particular *combination* of regulatory proteins, some of which are lacking in cells that do not respond to MyoD. The determination and differentiation of other kinds of tissues may play out in a similar fashion. A growing body of experimental evidence supports the idea that master regulatory proteins like MyoD might actually function by opening the chromatin in particular regions. This allows access to transcription machinery for activation of the next set of cell-type-specific genes.

What is the molecular basis for muscle cell differentiation? The MyoD protein is a transcription factor (see Figure 18.9) that binds to specific control elements in the enhancers of various target genes and stimulates their expression **(Figure 18.18)**. Some target genes for MyoD encode still other muscle-specific transcription factors. MyoD also stimulates expression of the

Figure 18.18 Determination and differentiation of muscle cells. Skeletal muscle cells arise from embryonic cells as a result of changes in gene expression. (In this depiction, the process of gene activation is greatly simplified.)



1 Determination. Signals from other cells lead to activation of a master regulatory gene called *myoD*, and the cell makes MyoD protein, a specific transcription factor that acts as an activator. The cell, now called a myoblast, is irreversibly committed to becoming a skeletal muscle cell.

2 Differentiation. MyoD protein stimulates the *myoD* gene further and activates genes encoding other muscle-specific transcription factors, which in turn activate genes for muscle proteins. MyoD also turns on genes that block the cell cycle, thus stopping cell division. The nondividing myoblasts fuse to become mature multinucleate muscle cells, also called muscle fibers.

WHAT IF? ➤ *What would happen if a mutation in the* myoD *gene resulted in the production of an altered MyoD protein that could not activate the* myoD *gene?*

myoD gene itself, an example of positive feedback that perpetuates MyoD's effect in maintaining the cell's differentiated state. Presumably, all the genes activated by MyoD have enhancer control elements recognized by MyoD and are thus coordinately controlled. Finally, the secondary transcription factors activate the genes for proteins such as myosin and actin that confer the unique properties of skeletal muscle cells.

We have now seen how different programs of gene expression that are activated in the fertilized egg can result in differentiated cells and tissues. But for the tissues to function effectively in the organism as a whole, the organism's *body plan*—its overall three-dimensional arrangement—must be established and superimposed on the differentiation process. Next we'll investigate the molecular basis for the establishment of the body plan, using the well-studied fruit fly *Drosophila melanogaster* as an example.

Pattern Formation: Setting Up the Body Plan

Cytoplasmic determinants and inductive signals both contribute to the development of a spatial organization in which the tissues and organs of an organism are all in their characteristic places. This process is referred to as **pattern formation**.

Just as the locations of the front, back, and sides of a new building are determined before construction begins, pattern formation in animals begins in the early embryo, when the major axes of an animal are established. In a bilaterally symmetrical animal, the relative positions of head and tail, right and left sides, and back and front—the three major body axes—are set up before the organs appear. The molecular cues that control pattern formation, collectively called **positional information**, are provided by cytoplasmic determinants and inductive signals (see Figure 18.17). These cues tell a cell its location relative to the body axes and to neighboring cells and determine how the cell and its descendants will respond to future molecular signals.

During the first half of the 20th century, classical embryologists made detailed anatomical observations of embryonic development in a number of species and performed experiments in which they manipulated embryonic tissues. Although this research laid the groundwork for understanding the mechanisms of development, it did not reveal the

specific molecules that guide development or determine how patterns are established.

In the 1940s, scientists began using the genetic approach—the study of mutants—to investigate *Drosophila* development. That approach has had spectacular success. These studies have established that genes control development and have led to an understanding of the key roles that specific molecules play in defining position and directing differentiation. By combining anatomical, genetic, and biochemical approaches to the study of *Drosophila* development, researchers have discovered developmental principles common to many other species, including humans.

The Life Cycle of Drosophila

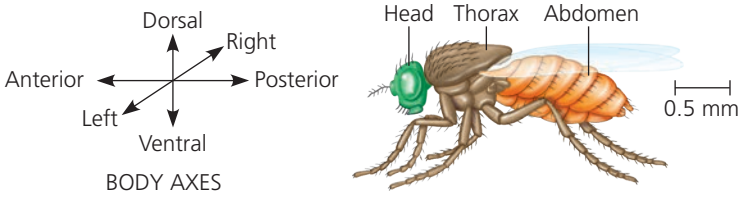
Fruit flies and other arthropods have a modular construction, an ordered series of segments. These segments make up the body's three major parts: the head, the thorax (the mid-body, from which the wings and legs extend), and the abdomen **(Figure 18.19a)**. Like other bilaterally symmetrical animals, *Drosophila* has an anterior-posterior (head-to-tail) axis, a dorsal-ventral (back-to-belly) axis, and a right-left axis. In *Drosophila*, cytoplasmic determinants that are localized in the unfertilized egg provide positional information for the placement of anterior-posterior and dorsal-ventral axes even before fertilization. We'll focus here on the molecules involved in establishing the anterior-posterior axis.

The *Drosophila* egg develops in one of the female's ovaries, adjacent to ovarian cells called nurse cells and surrounded by so-called follicle cells (**Figure 18.19b**, top). These support cells supply the egg with nutrients, mRNAs, and other substances needed for development and make the egg shell. After fertilization and laying of the egg, embryonic development results in the formation of a segmented larva, which goes through three larval stages. Then, in a process much like that by which a caterpillar becomes a butterfly, the fly larva forms a pupa in which it metamorphoses into the adult fly pictured in Figure 18.19a.

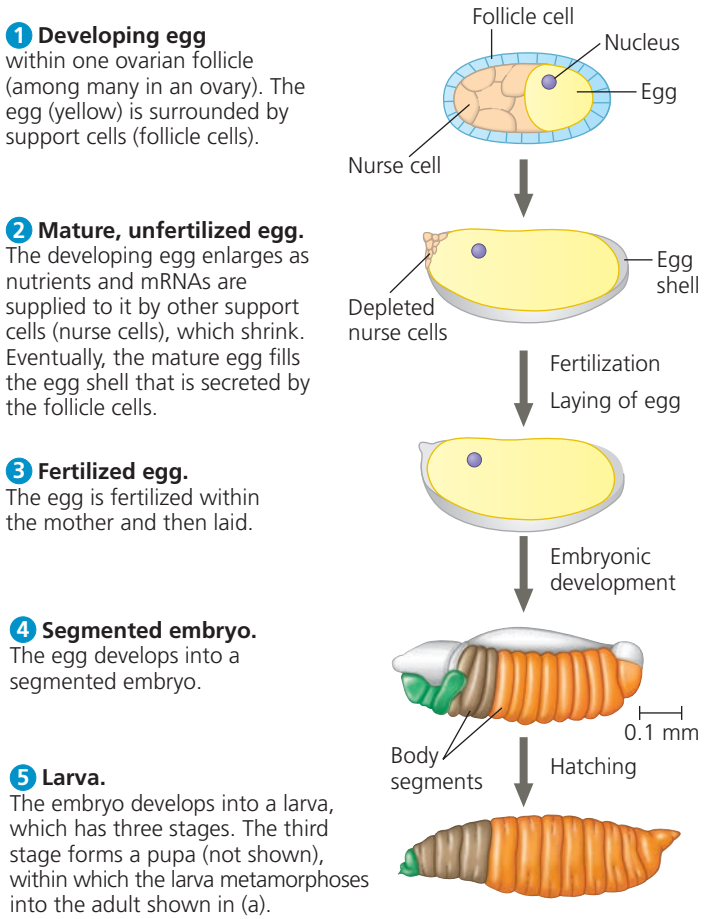
Genetic Analysis of Early Development: Scientific Inquiry

Edward B. Lewis was a visionary American biologist who, in the 1940s, first showed the value of the genetic approach to studying embryonic development in *Drosophila*. Lewis studied bizarre mutant flies with developmental defects that led to extra wings or legs in the wrong place **(Figure 18.20)**. He located the mutations on the fly's genetic map, thus connecting the developmental abnormalities to specific genes. This research supplied the first concrete evidence that genes somehow direct the developmental processes studied by embryologists. The genes Lewis discovered, called **homeotic genes**, are regulatory genes that control pattern formation in the late embryo, larva, and adult.

Figure 18.19 Key events in the *Drosophila* life cycle.

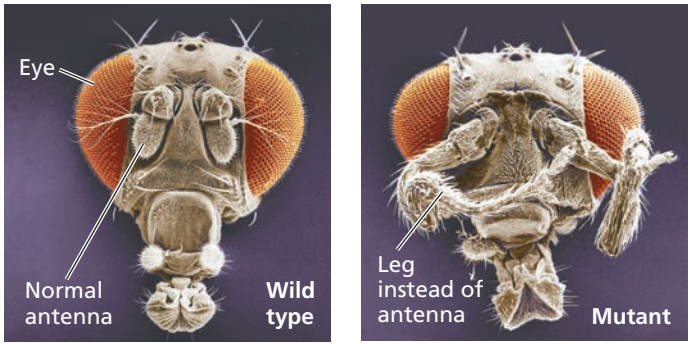


(a) Adult. The adult fly is segmented, and multiple segments make up each of the three main body parts—head, thorax, and abdomen. The body axes are shown by arrows.



(b) Development from egg to larva.

Figure 18.20 Abnormal pattern formation in *Drosophila*. Mutations in homeotic genes cause abnormal placement of structures in an animal, such as the legs extending from the mutant fly's head in place of antenna. (Colorized SEMs)



Further insight into pattern formation during early embryonic development did not come for another 30 years, when two researchers in Germany, Christiane Nüsslein-Volhard and Eric Wieschaus, set out to identify *all* the genes that affect segment formation in *Drosophila*. The project was daunting for three reasons. The first was the sheer number of *Drosophila* genes, now known to total about 14,000. The genes affecting segmentation might be just a few needles in a haystack or might be so numerous and varied that the scientists would be unable to make sense of them. Second, mutations affecting a process as fundamental as segmentation would surely be **embryonic lethals**, mutations with phenotypes causing death at the embryonic or larval stage. Because organisms with embryonic lethal mutations never reproduce, they cannot be bred for study. The researchers dealt with this problem by looking for recessive mutations, which can be propagated in heterozygous flies that act as genetic carriers. Third, cytoplasmic determinants in the egg were known to play a role in axis formation, so the researchers knew they would have to study the mother's genes as well as those of the embryo. It is the mother's genes that we will discuss further as we focus on how the anterior-posterior body axis is set up in the developing egg.

Nüsslein-Volhard and Wieschaus began their search for segmentation genes by exposing flies to a mutagenic chemical that affected the flies' gametes. They mated the mutagenized flies and then scanned their descendants for dead embryos or larvae with abnormal segmentation or other defects. For example, to find genes that might set up the anterior-posterior axis, they looked for embryos or larvae with abnormal ends, such as two heads or two tails, predicting that such abnormalities would arise from mutations in maternal genes required for correctly setting up the offspring's head or tail end.

Using this approach, Nüsslein-Volhard and Wieschaus eventually identified about 1,200 genes essential for pattern formation during embryonic development. Of these, about 120 were essential for normal segmentation. Over several years, the researchers were able to group these segmentation genes by general function, to find them on the fly's chromosomes, and to isolate many of them for further study in the lab. The result was a detailed molecular understanding of the early steps in pattern formation in *Drosophila*.

When the results of Nüsslein-Volhard and Wieschaus were combined with Lewis's earlier work, a coherent picture of *Drosophila* development emerged. In recognition of their discoveries, the three researchers were awarded a Nobel Prize in 1995. Next, let's consider a specific example of the genes that Nüsslein-Volhard, Wieschaus, and co-workers found.

Axis Establishment

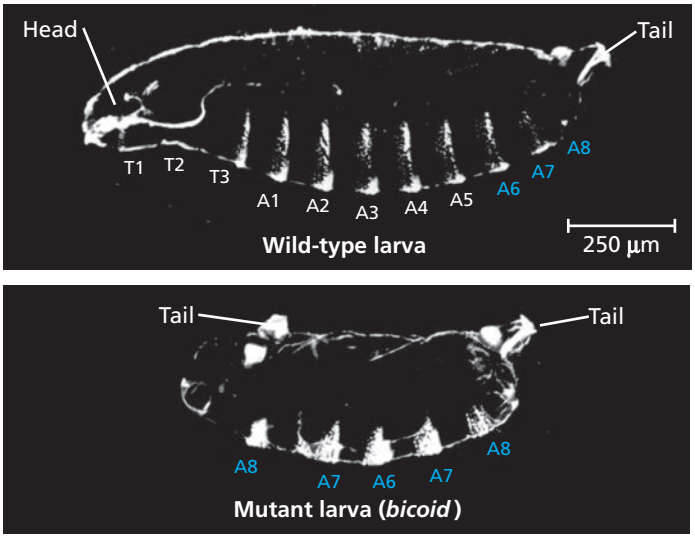
As we mentioned earlier, cytoplasmic determinants in the egg are the substances that initially establish the axes of the *Drosophila* body. These substances are encoded by genes of the mother, fittingly called maternal effect genes. A gene classified as a **maternal effect gene** is one that, when mutant in the mother, results in a mutant phenotype in the offspring, regardless of the offspring's own genotype. In fruit fly development, the mRNA or protein products of maternal effect genes are placed in the egg while it is still in the mother's ovary. When the mother has a mutation in such a gene, she makes a defective gene product (or none at all), and her eggs are defective; when these eggs are fertilized, they fail to develop properly.

Because they control the orientation (polarity) of the egg and consequently that of the fly, these maternal effect genes are also called *egg-polarity genes*. One group of these genes sets up the anterior-posterior axis of the embryo, while a second group establishes the dorsal-ventral axis. Like mutations in segmentation genes, mutations in maternal effect genes are generally embryonic lethals.

Bicoid: A Morphogen That Determines Head Structures

To see how maternal effect genes determine the body axes of the offspring, we will focus on one such gene, called ***bicoid***, a term meaning "two-tailed." An embryo or larva whose mother has two mutant *bicoid* alleles lacks the front half of its body and has posterior structures at both ends **(Figure 18.21)**. This phenotype suggested to Nüsslein-Volhard and her colleagues that the product of the mother's *bicoid* gene is essential for setting up the anterior end of the fly and might be concentrated at the future anterior end of the embryo. This hypothesis is an example of the *morphogen gradient hypothesis* first proposed

Figure 18.21 Effect of the *bicoid* gene on *Drosophila* development. A wild-type fruit fly larva has a head, three thoracic (T) segments, eight abdominal (A) segments, and a tail. A larva whose mother has two mutant alleles of the *bicoid* gene has two tails and lacks all anterior structures (LMs).



by embryologists a century ago, in which gradients of substances called **morphogens** establish an embryo's axes and other features of its form.

DNA technology and other modern biochemical methods enabled the researchers to test whether the *bicoid* product, a protein called Bicoid, is in fact a morphogen that determines the anterior end of the fly. The first question they asked was whether the mRNA and protein products of this gene are located in the egg in a position consistent with the hypothesis. They found that *bicoid* mRNA is highly concentrated at the extreme anterior end of the mature egg **(Figure 18.22)**. After the egg is fertilized, the mRNA is translated into protein. The Bicoid protein then diffuses from the anterior end toward the posterior, resulting in a gradient of protein within the early embryo, with the highest concentration at the anterior end. These results are consistent with the hypothesis that Bicoid protein specifies the fly's anterior end. To test the hypothesis more specifically, scientists injected pure *bicoid* mRNA into various regions of early embryos. The protein that resulted from its translation caused anterior structures to form at the injection sites.

The *bicoid* research was groundbreaking for several reasons. First, it led to the identification of a specific protein required for some of the earliest steps in pattern formation. It thus helped us understand how different regions of the egg can give rise to cells that go down different developmental pathways. Second, it increased our understanding of the mother's critical role in the initial phases of embryonic development. Finally, the principle that a gradient of morphogens can determine polarity and position has proved to be a key developmental concept for a number of species, just as early embryologists had hypothesized.

Maternal mRNAs are crucial during development of many species. In *Drosophila*, gradients of specific proteins encoded by maternal mRNAs not only determine the posterior and anterior ends but also establish the dorsal-ventral axis. As the fly embryo grows, it reaches a point when the embryonic program of gene expression takes over, and the maternal mRNAs must be destroyed. (This process involves miRNAs in *Drosophila* and other species.) Later, positional information encoded by the embryo's genes, operating on an ever finer scale, establishes a specific number of correctly oriented segments and triggers the formation of each segment's characteristic structures. When the genes operating in this final step are abnormal, the pattern of the adult is abnormal, as you saw in Figure 18.20.

Interview with Nancy Hopkins: Studying the genetic basis of development in zebrafish

Evolutionary Developmental Biology ("Evo-Devo")

EVOLUTION The fly with legs emerging from its head in Figure 18.20 is the result of a single mutation in one gene, a homeotic gene. The gene does not encode any antenna protein, however. Instead, it encodes a transcription factor that regulates other genes, and its malfunction leads to misplaced structures, such as legs instead of antennae. The observation that a change in gene regulation during development could lead to such a fantastic change in body form prompted some scientists to consider whether these types of mutations could contribute to evolution by generating novel body shapes. Ultimately, this line of inquiry gave rise to the field of evolutionary developmental biology, so-called "evo-devo," which will be further discussed in Concept 20.6.

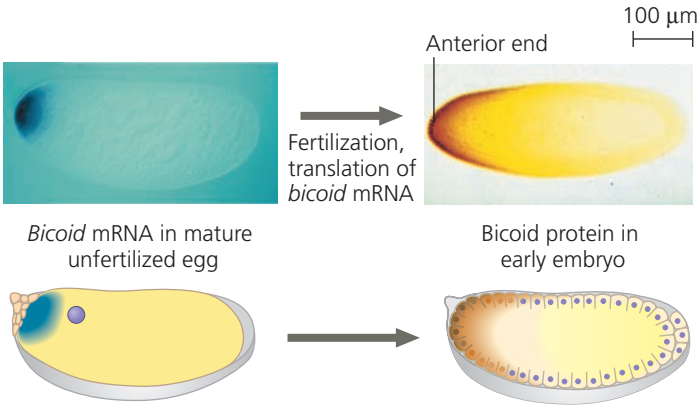
In this section, we have seen how a carefully orchestrated program of sequential gene regulation controls the

Figure 18.22

Inquiry Could Bicoid be a morphogen that determines the anterior end of a fruit fly?

Experiment Using a genetic approach to study *Drosophila* development, Christiane Nüsslein-Volhard and colleagues at two research institutions in Germany analyzed expression of the *bicoid* gene. The researchers hypothesized that *bicoid* normally codes for a morphogen that specifies the head (anterior) end of the embryo. To begin to test this hypothesis, they used molecular techniques to determine whether the mRNA and protein encoded by this gene were found in the anterior end of the fertilized egg and early embryo of wild-type flies.

Results *Bicoid* mRNA (dark blue in the light micrographs and drawings) was confined to the anterior end of the unfertilized egg. Later in development, Bicoid protein (dark orange) was seen to be concentrated in cells at the anterior end of the embryo.



Conclusion The location of *bicoid* mRNA and the diffuse gradient of Bicoid protein seen later are consistent with the hypothesis that Bicoid protein is a morphogen specifying formation of head-specific structures.

Further Reading C. Nüsslein-Volhard et al., Determination of anteroposterior polarity in *Drosophila*, *Science* 238:1675–1681 (1987); W. Driever and C. Nüsslein-Volhard, A gradient of Bicoid protein in *Drosophila* embryos, *Cell* 54:83–93 (1988); T. Berleth et al., The role of localization of *bicoid* RNA in organizing the anterior pattern of the *Drosophila* embryo, *EMBO Journal* 7:1749–1756 (1988).

WHAT IF? ➤ *The researchers needed further evidence, so they injected bicoid mRNA into the anterior end of an egg from a female with a mutation disabling the bicoid gene. Given that the hypothesis was supported, what must their results have been?*

Animation: Role of *bicoid* Gene in *Drosophila* Development

transformation of a fertilized egg into a multicellular organism. The program is carefully balanced between turning on the genes for differentiation in the right place and turning off other genes. Even when an organism is fully developed, gene expression is regulated in a similarly fine-tuned manner. In the final section of the chapter, we'll consider how fine this tuning is by looking at how specific changes in expression of just a few genes can lead to the development of cancer.

CONCEPT CHECK 18.4

1. **MAKE CONNECTIONS ➤** As you learned in Chapter 12, mitosis gives rise to two daughter cells that are genetically identical to the parent cell. Yet you, the product of many mitotic divisions, are not composed of identical, zygote-like cells. Why?
2. **MAKE CONNECTIONS ➤** Explain how the signaling molecules released by an embryonic cell can induce changes in a neighboring cell without entering the cell. (See Figures 9.15 and 9.16.)
3. How do fruit fly maternal effect genes determine the polarity of the egg and the embryo?
4. **WHAT IF? ➤** In Figure 18.17b, the lower cell is synthesizing signaling molecules, whereas the upper cell is expressing receptors for these molecules. In terms of gene regulation and cytoplasmic determinants, explain how these cells came to synthesize different molecules.

For suggested answers, see Appendix A.

CONCEPT 18.5

Cancer results from genetic changes that affect cell cycle control

In Concept 12.3, we considered cancer as a type of disease in which cells escape from the control mechanisms that normally limit their growth. Now that we have discussed the molecular basis of gene expression and its regulation, we are ready to look at cancer more closely. The gene regulation systems that go wrong during cancer turn out to be the very same systems that play important roles in embryonic development, the immune response, and many other biological processes. Thus, research into the molecular basis of cancer has both benefited from and informed many other fields of biology.

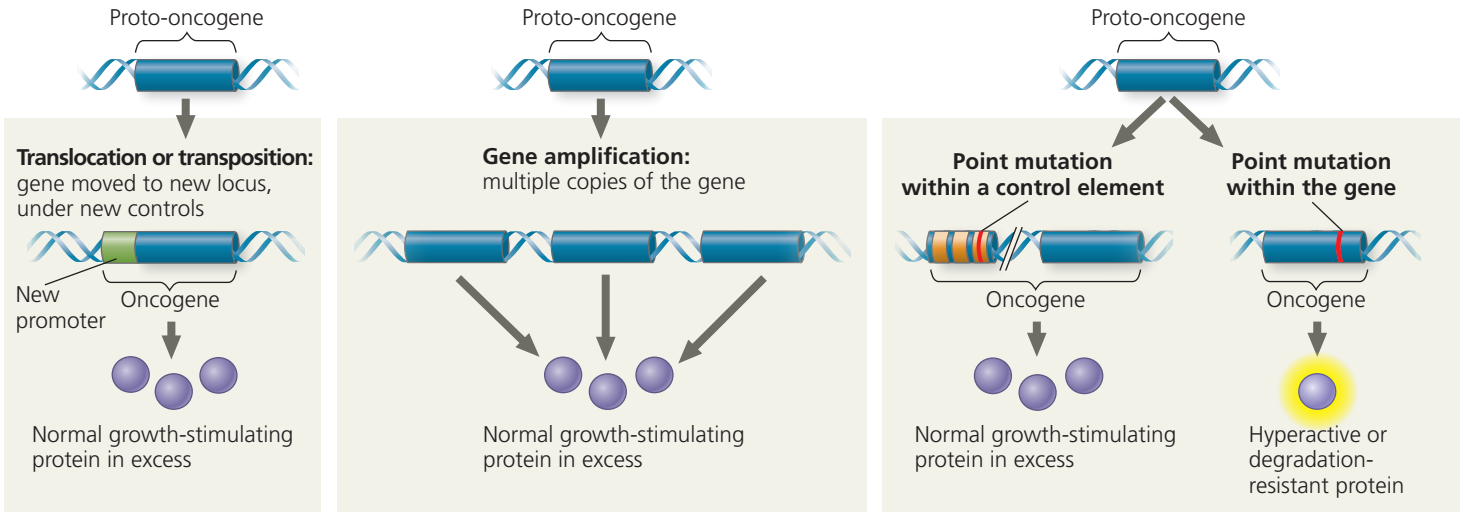
Types of Genes Associated with Cancer

The genes that normally regulate cell growth and division during the cell cycle include genes for growth factors, their receptors, and the intracellular molecules of signaling pathways. (To review cell signaling, see Concept 9.2; for regulation of the cell cycle, see Concept 12.3.) Mutations that alter any of these genes in somatic cells can lead to cancer. The agent of such change can be random spontaneous mutation. However, it is also likely that many cancer-causing mutations result from environmental influences, such as chemical carcinogens, X-rays and other high-energy radiation, and some viruses.

Cancer research led to the discovery of cancer-causing genes called **oncogenes** (from the Greek *onco*, tumor) in certain types of viruses. Subsequently, close counterparts of viral oncogenes were found in the genomes of humans and other animals. The normal versions of the cellular genes, called **proto-oncogenes**, code for proteins that stimulate normal cell growth and division.

How might a proto-oncogene—a gene that has an essential function in normal cells—become an oncogene, a cancer-causing gene? In general, an oncogene arises from a genetic change that leads to an increase either in the amount of the proto-oncogene's protein product or in the intrinsic activity of each protein molecule. The genetic changes that convert proto-oncogenes to oncogenes fall into three main categories: movement of DNA within the genome, amplification of a proto-oncogene, and point mutations in a control element or in the proto-oncogene itself **(Figure 18.23)**.

Figure 18.23 Genetic changes that can turn proto-oncogenes into oncogenes.



Cancer cells are frequently found to contain chromosomes that have broken and rejoined incorrectly, translocating fragments from one chromosome to another (see Figure 15.14). Having learned how gene expression is regulated, you can now see the possible consequences of such translocations. If a translocated proto-oncogene ends up near an especially active promoter (or other control element), its transcription may increase, making it an oncogene. The second main type of genetic change, amplification, increases the number of copies of the proto-oncogene in the cell through repeated gene duplication (discussed in Concept 20.5). The third possibility is a point mutation either in the promoter or an enhancer that controls a proto-oncogene, causing an increase in its expression, or in the coding sequence of the proto-oncogene, changing the gene's product to a protein that is more active or more resistant to degradation than the normal protein. These mechanisms can lead to abnormal stimulation of the cell cycle and put the cell on the path to becoming a cancer cell.

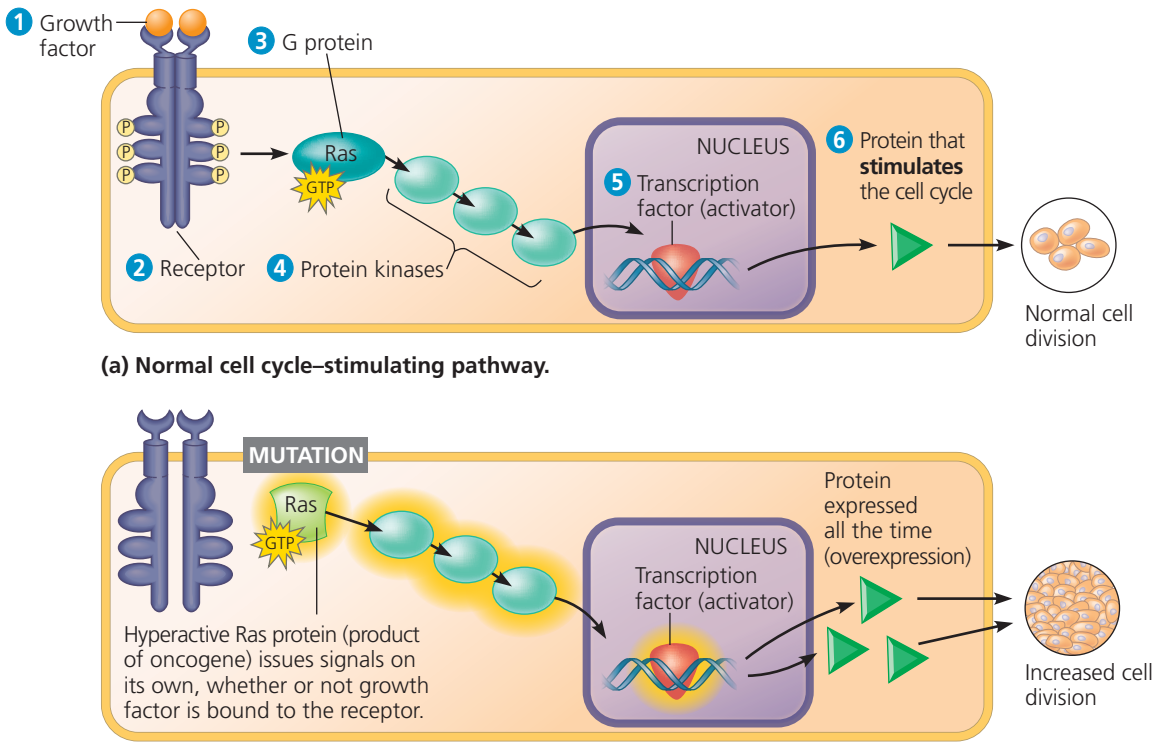
In addition to genes whose products normally promote cell division, cells contain genes whose normal products *inhibit* cell division. Such genes are called **tumor-suppressor genes** since the proteins they encode help prevent uncontrolled cell growth. Any mutation that decreases the normal activity of a tumor-suppressor protein may contribute to the onset of cancer, in effect stimulating growth through the absence of suppression.

The protein products of tumor-suppressor genes have various functions. Some repair damaged DNA, a function that prevents the cell from accumulating cancer-causing mutations. Other tumor-suppressor proteins control the adhesion of cells to each other or to the extracellular matrix; proper cell anchorage is crucial in normal tissues—and is often absent in cancers. Still other tumor-suppressor proteins are components of cell-signaling pathways that inhibit the cell cycle.

Interference with Normal Cell-Signaling Pathways

The proteins encoded by many proto-oncogenes and tumor-suppressor genes are components of cell-signaling pathways. Let's take a closer look at how such proteins function in normal cells and what goes wrong with their function in cancer cells. We will focus on the products of two key genes, the *ras* proto-oncogene and the *p53* tumor-suppressor gene. Mutations in *ras* occur in about 30% of human cancers, and mutations in *p53* in more than 50%.

The Ras protein, encoded by the ***ras* gene** (named for rat sarcoma, a connective tissue cancer), is a G protein that relays a signal from a growth factor receptor on the plasma membrane to a cascade of protein kinases (see Figures 9.8 and 9.10). The cellular response at the end of the pathway is the synthesis of a protein that stimulates the cell cycle **(Figure 18.24a)**. Normally, such a pathway will not operate unless triggered by the appropriate growth factor. But certain mutations in the *ras* gene can lead to production of a hyperactive Ras protein that triggers the kinase cascade even in



(a) Normal cell cycle–stimulating pathway.

(b) Mutant cell cycle–stimulating pathway.

➤ Figure 18.24 Normal and mutant cell cycle–stimulating pathway. **(a)** The normal pathway is triggered by 1 a growth factor that binds to 2 its receptor in the plasma membrane. The signal is relayed to 3 a G protein called Ras. Like all G proteins, Ras is active when GTP is bound to it. Ras passes the signal to 4 a series of protein kinases. The last kinase activates 5 a transcription factor (activator) that turns on one or more genes for 6 a protein that stimulates the cell cycle. **(b)** If a mutation makes Ras or any other pathway component abnormally active, excessive cell division and cancer may result.

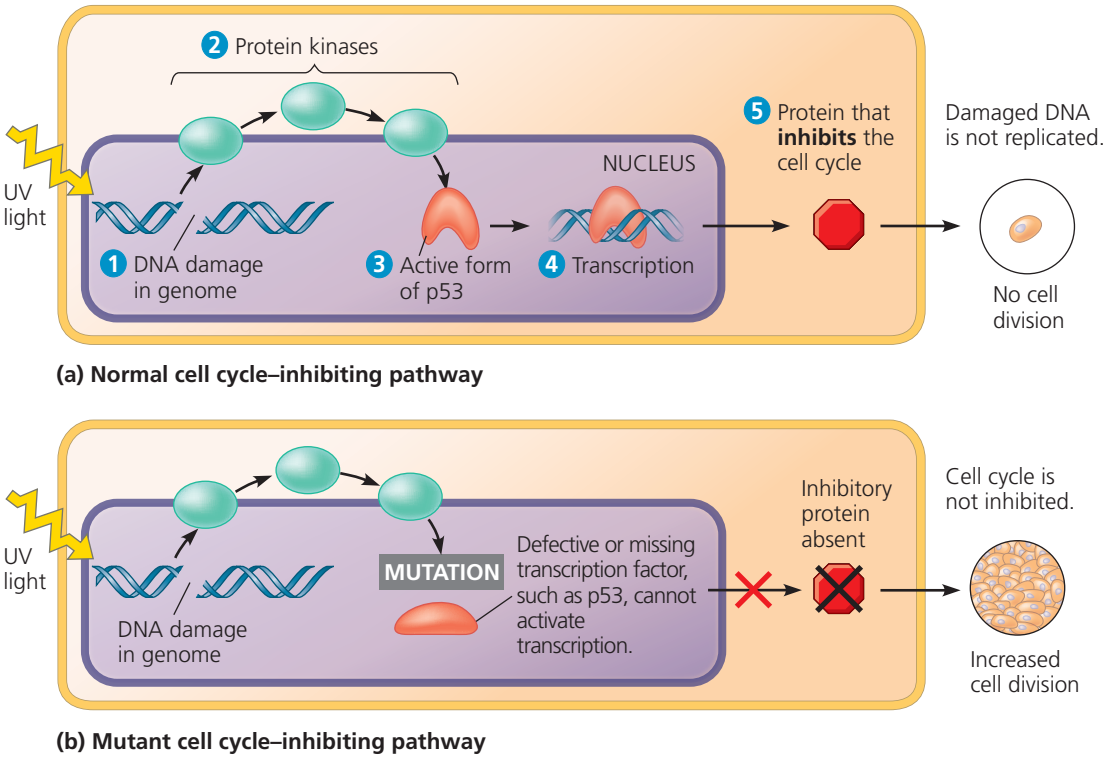
the absence of growth factor, resulting in increased cell division **(Figure 18.24b)**. In fact, hyperactive versions or excess amounts of any of the pathway's components can have the same outcome: excessive cell division.

Figure 18.25a shows a pathway in which an intracellular signal leads to the synthesis of a protein that suppresses the cell cycle. In this case, the signal is damage to the cell's DNA, perhaps as the result of exposure to ultraviolet light. Operation of this signaling pathway blocks the cell cycle until the damage has been repaired. Otherwise, the damage might contribute to tumor formation by causing mutations or chromosomal abnormalities. Thus, the genes for the components of the pathway act as tumor-suppressor genes. The ***p53* gene**, named for the 53,000-dalton molecular weight of its protein product, is a tumor-suppressor gene. The protein it encodes is a specific transcription factor that promotes the synthesis of cell cycle–inhibiting proteins. That is why a mutation that knocks out the *p53* gene, like a mutation that leads to a hyperactive Ras protein, can lead to excessive cell growth and cancer **(Figure 18.25b)**.

The *p53* gene has been called the "guardian angel of the genome." Once the gene is activated—for example, by DNA damage—the p53 protein functions as an activator for several other genes. Often it activates a gene called *p21*, whose product halts the cell cycle by binding to cyclin-dependent kinases, allowing time for the cell to repair the DNA. Researchers recently showed that p53 also activates expression of a group of miRNAs, which in turn inhibit the cell cycle. In addition, the p53 protein can turn on genes directly involved in DNA repair. Finally, when DNA damage is irreparable, p53 activates "suicide" genes, whose protein products bring about programmed cell death (apoptosis; see Figure 9.20). Thus, p53 acts in several ways to prevent a cell from passing on mutations due to DNA damage. If mutations do accumulate and the cell survives through many divisions—as is more likely if the *p53* tumor-suppressor gene is defective or missing—cancer may ensue. The many functions of p53 suggest a complex picture of regulation in normal cells, one that we do not yet fully understand.

A recent study may underscore the protective role of *p53* while illuminating a long-standing research question: Why is cancer so rare among elephants? The incidence of cancer among elephants in zoo-based studies has been estimated at about 3%, compared to closer to 30% for humans. Genome sequencing revealed that elephants have 20 copies of the *p53* gene, compared to one copy in humans, other mammals, and even manatees, elephants' closest living relatives. There are undoubtedly other underlying reasons, but the correlation between low cancer rate and extra copies of the *p53* gene bears further investigation.

For the present, the diagrams in Figure 18.24 and Figure 18.25 are an accurate view of how mutations can contribute to cancer, but we still don't know exactly how a particular cell becomes a cancer cell. As we discover previously unknown



➤ Figure 18.25 Normal and mutant cell cycle–inhibiting pathway. (a) In the normal pathway, 1 DNA damage is an intracellular signal that is passed via 2 protein kinases, leading to activation of 3 p53. Activated p53 promotes 4 transcription of the gene for 5 a protein that inhibits the cell cycle. The resulting suppression of cell division ensures that the damaged DNA is not replicated. If the DNA damage is irreparable, then the p53 signal leads to programmed cell death (apoptosis). **(b)** Mutations causing deficiencies in any pathway component can contribute to the development of cancer.

? *Explain whether a cancer-causing mutation in a tumor-suppressor gene, such as* p53, *is more likely to be a recessive or a dominant mutation.*

aspects of gene regulation, it is informative to study their role in the onset of cancer. Such studies have shown, for instance, that DNA methylation and histone modification patterns differ in normal and cancer cells and that miRNAs probably participate in cancer development. While we've learned a lot about cancer by studying cell-signaling pathways, there is still much more to learn.

The Multistep Model of Cancer Development

More than one somatic mutation is generally needed to produce all the changes characteristic of a full-fledged cancer cell. This may help explain why the incidence of cancer increases greatly with age. If cancer results from an accumulation of mutations and if mutations occur throughout life, then the longer we live, the more likely we are to develop cancer.

The model of a multistep path to cancer is well supported by studies of one of the best-understood types of human cancer: colorectal cancer, which affects the colon and/or rectum. About 140,000 new cases of colorectal cancer are diagnosed each year in the United States, and the disease causes 50,000 deaths per year. Like most cancers, colorectal cancer develops gradually **(Figure 18.26)**. The first sign is often a polyp, a small, benign growth in the colon lining. The cells of the polyp look normal, although they divide unusually frequently. The tumor grows and may eventually become malignant, invading other tissues. The development of a malignant tumor is paralleled by a gradual accumulation of mutations that convert proto-oncogenes to oncogenes and knock out tumor-suppressor genes. A *ras* oncogene and a mutated *p53* tumor-suppressor gene are often involved.

About half a dozen changes must occur at the DNA level for a cell to become fully cancerous. These changes usually include the appearance of at least one active oncogene and the mutation or loss of several tumor-suppressor genes. Furthermore, since mutant tumor-suppressor alleles are usually recessive, in most cases mutations must knock out *both* alleles in a cell's genome to block tumor suppression. (Most oncogenes, on the other hand, behave as dominant alleles.)

Since we understand the progression of this type of cancer, routine screenings (colonoscopies, for example) are recommended to identify and remove any suspicious polyps. The colorectal cancer mortality rate has been declining for the past 20 years due to increased screening and improved treatments. Treatments for other cancers have improved as well. Advances in the sequencing of DNA and mRNA allow medical researchers to compare the genes expressed by different types of tumors and by the same type in different people. These comparisons have led to personalized treatments based on the molecular characteristics of a person's tumor.

Breast cancer is the second most common form of cancer in the United States, and the first among women. Each year, this cancer strikes over 230,000 women (and some men) in the United States and kills 40,000 (450,000 worldwide). A major problem with understanding breast cancer is its heterogeneity: Tumors differ in significant ways. Identifying differences between types of breast cancer is expected to improve treatment and decrease the mortality rate. In 2012, The Cancer Genome Atlas Network, sponsored by the National Institutes of Health, published the results of a multi-team effort that used a genomics approach to profile subtypes of breast cancer based on their molecular signatures. Four major types of breast cancer were identified **(Figure 18.27)**. It is now routine to screen for the presence of particular signaling receptors in any breast cancer tumors, and individuals with breast cancer, along with their physicians, can now make more informed decisions about their treatments.

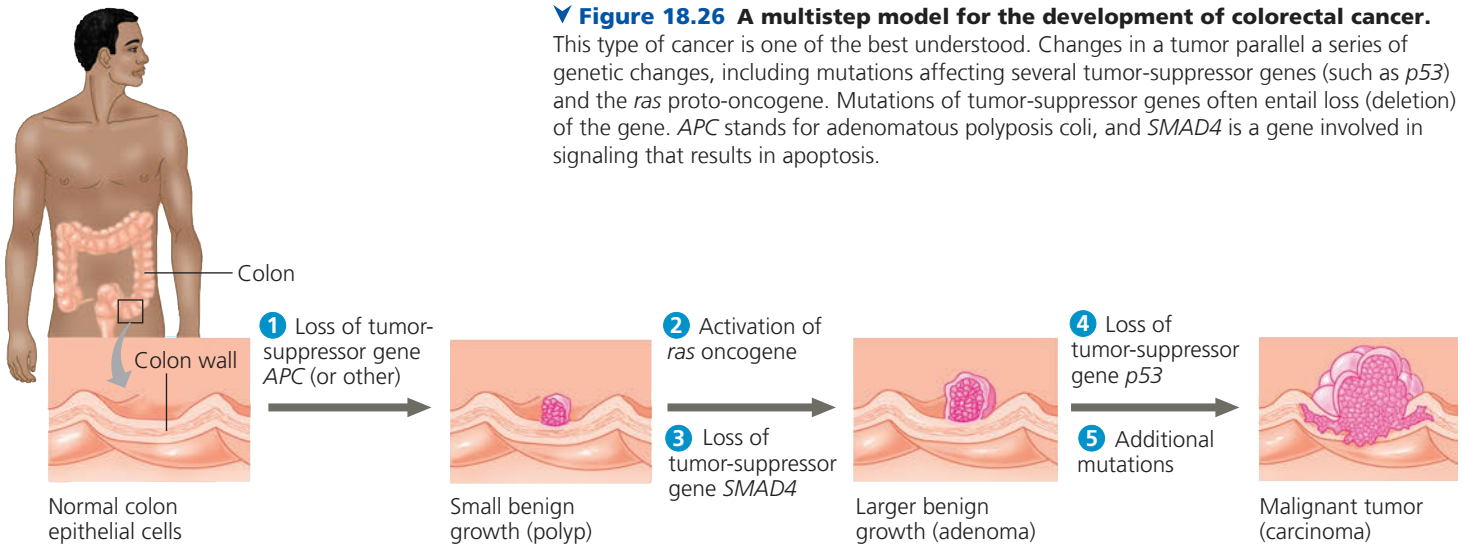


Figure 18.26 A multistep model for the development of colorectal cancer. This type of cancer is one of the best understood. Changes in a tumor parallel a series of genetic changes, including mutations affecting several tumor-suppressor genes (such as *p53*) and the *ras* proto-oncogene. Mutations of tumor-suppressor genes often entail loss (deletion) of the gene. *APC* stands for adenomatous polyposis coli, and *SMAD4* is a gene involved in signaling that results in apoptosis.

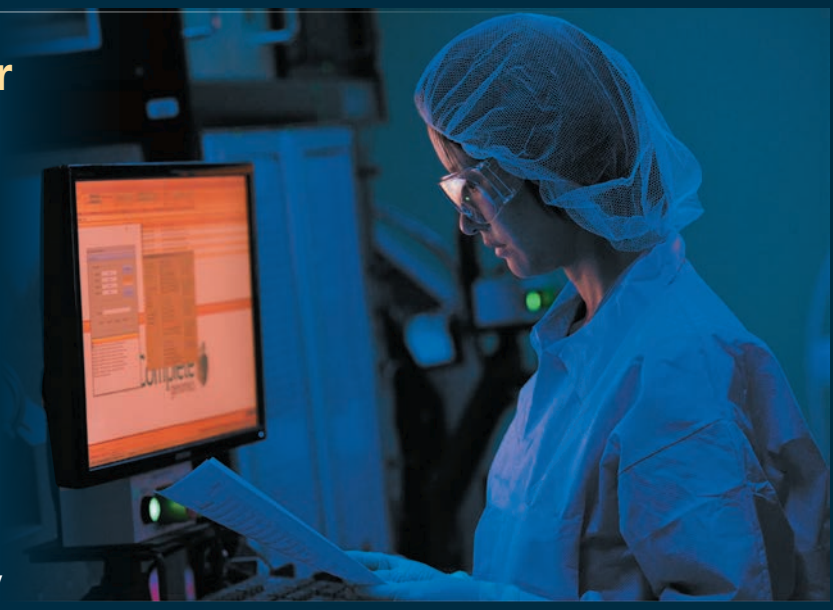
Figure 18.27 **MAKE CONNECTIONS**

Genomics, Cell Signaling, and Cancer

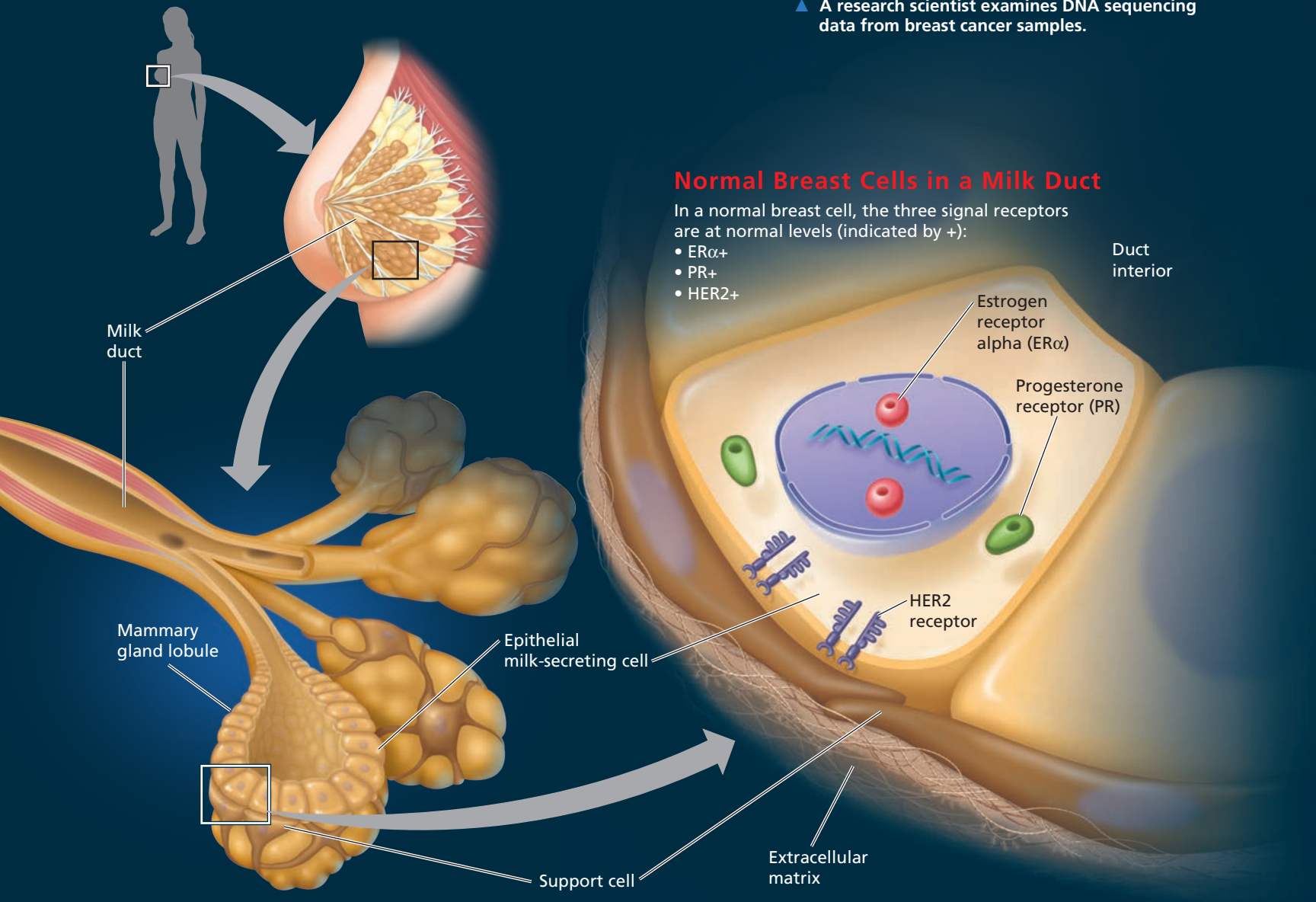
Modern medicine that melds genome-wide molecular studies with cell-signaling research is transforming the treatment of many diseases, such as breast cancer. Using microarray analysis (*see Figure 19.13*) and other techniques, researchers measured the relative levels of mRNA transcripts for every gene in hundreds of breast cancer tumor samples. They identified four major subtypes of breast cancer that differ in their expression of three signal receptors involved in regulating cell growth and division:

- Estrogen receptor alpha (ERα)
- Progesterone receptor (PR)
- HER2, a type of receptor called a receptor tyrosine kinase (*see Figure 9.8*)

(ERα and PR are steroid receptors; *see Figure 9.9*.) The absence or excess expression of these receptors can cause aberrant cell signaling, leading in some cases to inappropriate cell division, which may contribute to cancer (*see Figure 18.24*).



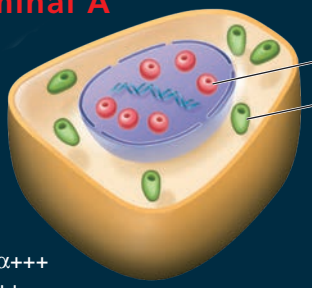
▲ **A research scientist examines DNA sequencing data from breast cancer samples.**



Breast Cancer Subtypes

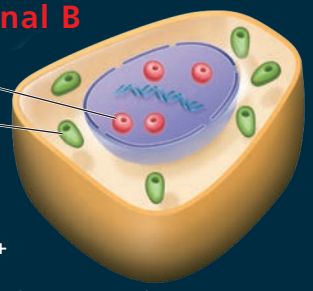
Each breast cancer subtype is characterized by the overexpression (indicated by ++ or +++) or absence (–) of three signal receptors: ERα, PR, and HER2. Breast cancer treatments are becoming more effective because they can be tailored to the specific cancer subtype.

Luminal A



- ERα+++
- PR++
- HER2–
- 40% of breast cancers
- Best prognosis

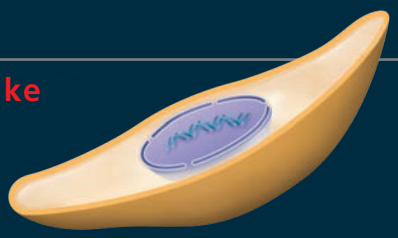
Luminal B



- ERα++
- PR++
- HER2– (shown here); some HER2++
- 15–20% of breast cancers
- Poorer prognosis than luminal A subtype

Both luminal subtypes overexpress ERα (luminal A more than luminal B) and PR, and usually lack expression of HER2. Both can be treated with drugs that target ERα and inactivate it, the most well-known drug being tamoxifen. These subtypes can also be treated with drugs that inhibit estrogen synthesis.

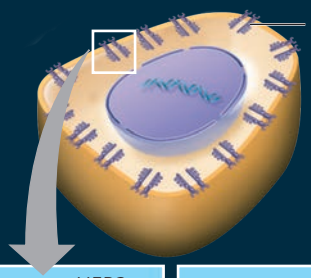
Basal-like



- ERα–
- PR–
- HER2–
- 15–20% of breast cancers
- More aggressive; poorer prognosis than other subtypes

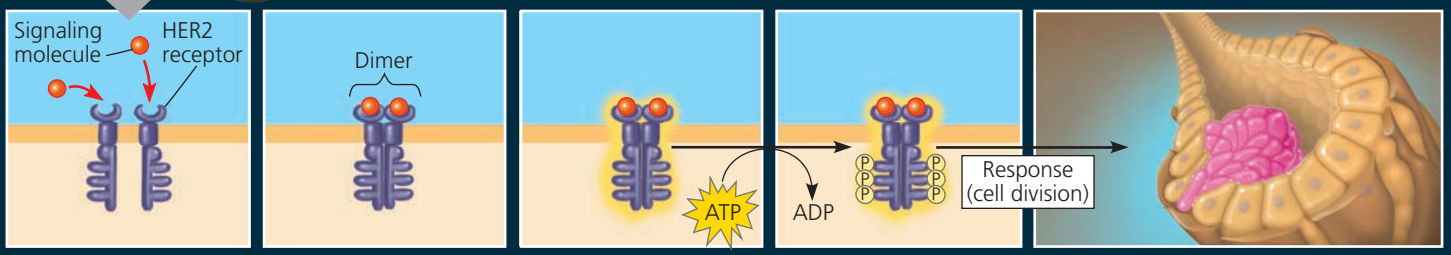
The basal-like subtype is "triple negative"—it does not express ERα, PR, or HER2. It often has a mutation in the tumor-suppressor gene *BRCA1* (*see Concept 18.5*). Treatments that target ER, PR, or HER2 are not effective, but new treatments are being developed. Currently, patients are treated with cytotoxic chemotherapy, which selectively kills fast-growing cells.

HER2



- ERα–
- PR–
- HER2++
- 10–15% of breast cancers
- Poorer prognosis than luminal A subtype

The HER2 subtype overexpresses HER2. Because it does not express either ERα or PR at normal levels, the cells are unresponsive to therapies that target those two receptors. However, patients with the HER2 subtype can be treated with Herceptin, an antibody protein that inactivates HER2 (*see Concept 12.3*).



1 Signaling molecules (such as a growth factor) bind to HER2 receptor monomers (single receptor proteins).

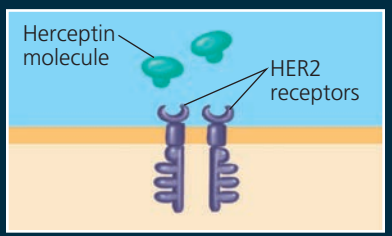
2 Binding of signaling molecules causes two receptor monomers to associate closely with each other, forming a dimer.

3 Formation of a dimer activates each monomer.

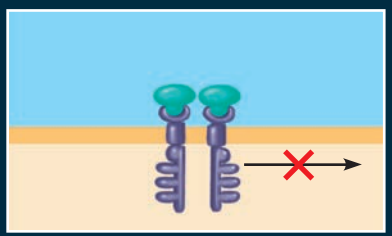
4 Each monomer adds phosphate from ATP to the other monomer, triggering a signal transduction pathway.

5 The signal is transduced through the cell, which leads to a cellular response—in this case, turning on genes that trigger cell division. HER2 cells have up to 100 times as many HER2 receptors as normal cells, so they undergo uncontrolled cell division.

Treatment with Herceptin for the HER2 subtype



1 The drug Herceptin binds to HER2 receptors in place of the usual signaling molecules.



2 In certain patients with the HER2 subtype, signaling is blocked and excessive cell division does not occur.

MAKE CONNECTIONS ➤ *When researchers compared gene expression in normal breast cells and cells from breast cancers, they found that the genes showing the most significant differences in expression encoded signal receptors, as shown here. Given what you learned in Chapters 9, 12, and this chapter, explain why this result is not surprising.*

Inherited Predisposition and Environmental Factors Contributing to Cancer

The fact that multiple genetic changes are required to produce a cancer cell helps explain the observation that cancers can run in families. An individual inheriting an oncogene or a mutant allele of a tumor-suppressor gene is one step closer to accumulating the necessary mutations for cancer to develop than is an individual without any such mutations.

Geneticists are working to identify inherited cancer alleles so that predisposition to certain cancers can be detected early in life. About 15% of colorectal cancers, for example, involve inherited mutations. Many of these affect the tumor-suppressor gene called *adenomatous polyposis coli*, or *APC* (see Figure 18.26). This gene has multiple functions in the cell, including regulation of cell migration and adhesion. Even in patients with no family history of the disease, the *APC* gene is mutated in 60% of colorectal cancers. In these individuals, new mutations must occur in both *APC* alleles before the gene's function is lost. Since only 15% of colorectal cancers are associated with known inherited mutations, researchers continue to try to identify "markers" that could predict the risk of developing this type of cancer.

Given the prevalence and significance of breast cancer, it is not surprising that it was one of the first cancers for which the role of inheritance was investigated. It turns out that for 5–10% of patients with breast cancer, there is evidence of a strong inherited predisposition. Geneticist Mary-Claire King began working on this problem in the mid-1970s. After 16 years of research, she convincingly demonstrated that mutations in one gene—*BRCA1*—were associated with increased susceptibility to breast cancer, a finding that flew in the face of medical opinion at the time. (*BRCA* stands for breast cancer.) Mutations in that gene or a gene called *BRCA2* are found in at least half of inherited breast cancers, and tests using DNA sequencing can detect these mutations. A woman who inherits one mutant *BRCA1* allele has a 60% probability of developing breast cancer before the age of 50, compared with only a 2% probability for an individual homozygous for the normal allele.

BRCA1 and *BRCA2* are considered tumor-suppressor genes because their wild-type alleles protect against breast cancer and their mutant alleles are recessive. (Note that mutations in *BRCA1* are commonly found in the genomes of cells from basal-like breast cancers; see Figure 18.27.) The BRCA1 and BRCA2 proteins both appear to function in the cell's DNA damage repair pathway. More is known about BRCA2, which, along with another protein, helps repair breaks that occur in both strands of DNA; this repair function is crucial for maintaining undamaged DNA.

MB **Interview with Mary-Claire King: Discovering breast cancer genes**

Because DNA breakage can contribute to cancer, it makes sense that the risk of cancer can be lowered by minimizing exposure to DNA-damaging agents, such as the ultraviolet radiation in sunlight and chemicals found in cigarette smoke. Novel genomics-based analyses of specific cancers, such as the approach described in Figure 18.27, are contributing to both early diagnosis and development of treatments that interfere with expression of key genes in tumors. Ultimately, such approaches are expected to lower the death rate from cancer.

The Role of Viruses in Cancer

The study of genes associated with cancer, inherited or not, increases our basic understanding of how disruption of normal gene regulation can result in this disease. In addition to the mutations and other genetic alterations described in this section, a number of *tumor viruses* can cause cancer in various animals, including humans. In fact, one of the earliest breakthroughs in understanding cancer came in 1911, when Peyton Rous, an American pathologist, discovered a virus that causes cancer in chickens. The Epstein-Barr virus, which causes infectious mononucleosis, has been linked to several types of cancer in humans, notably Burkitt's lymphoma. Papillomaviruses are associated with cancer of the cervix, and a virus called HTLV-1 causes a type of adult leukemia. Viruses play a role in about 15% of the cases of human cancer.

Viruses may at first seem very different from mutations as a cause of cancer. However, we now know that viruses can interfere with gene regulation in several ways if they integrate their genetic material into the DNA of a cell. Viral integration may donate an oncogene to the cell, disrupt a tumor-suppressor gene, or convert a proto-oncogene to an oncogene. Some viruses produce proteins that inactivate p53 and other tumor-suppressor proteins, making the cell more prone to becoming cancerous. Viruses are powerful biological agents; you'll learn more about them in Chapter 26.

MB **Animation: Causes of Cancer**

CONCEPT CHECK 18.5

1. Cancer-promoting mutations are likely to have different effects on the activity of proteins encoded by proto-oncogenes than they do on proteins encoded by tumor-suppressor genes. Explain.
2. Under what circumstances is cancer considered to have a hereditary component?
3. **MAKE CONNECTIONS ➤** The p53 protein can activate genes involved in apoptosis. Review Concept 9.5, and discuss how mutations in genes coding for proteins that function in apoptosis could contribute to cancer.

For suggested answers, see Appendix A.

18 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

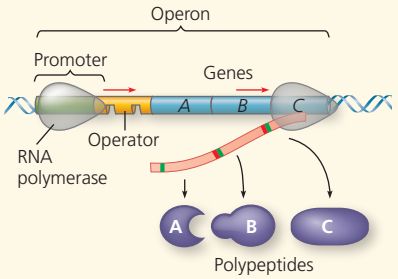
SUMMARY OF KEY CONCEPTS

CONCEPT 18.1

Bacteria often respond to environmental change by regulating transcription *(pp. 414–418)*

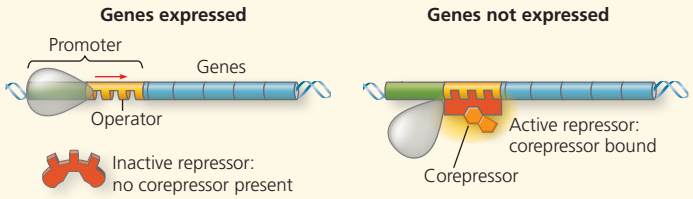
VOCAB SELF-QUIZ goo.gl/Rn5Uax

- Cells control metabolism by regulating enzyme activity or the expression of genes coding for enzymes. In bacteria, genes are often clustered into **operons**, with one promoter serving several adjacent genes. An **operator** site on the DNA switches the operon on or off, resulting in coordinate regulation of the genes.



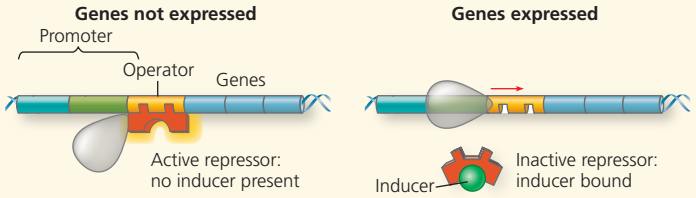
- Both repressible and inducible operons are examples of negative gene regulation. In either type of operon, binding of a specific **repressor** protein to the operator shuts off transcription. (The repressor is encoded by a separate **regulatory gene**.) In a repressible operon, the repressor is active when bound to a **corepressor**, usually the end product of an anabolic pathway.

Repressible operon:



- In an inducible operon, binding of an **inducer** to an innately active repressor inactivates the repressor and turns on transcription. Inducible enzymes usually function in catabolic pathways.

Inducible operon:

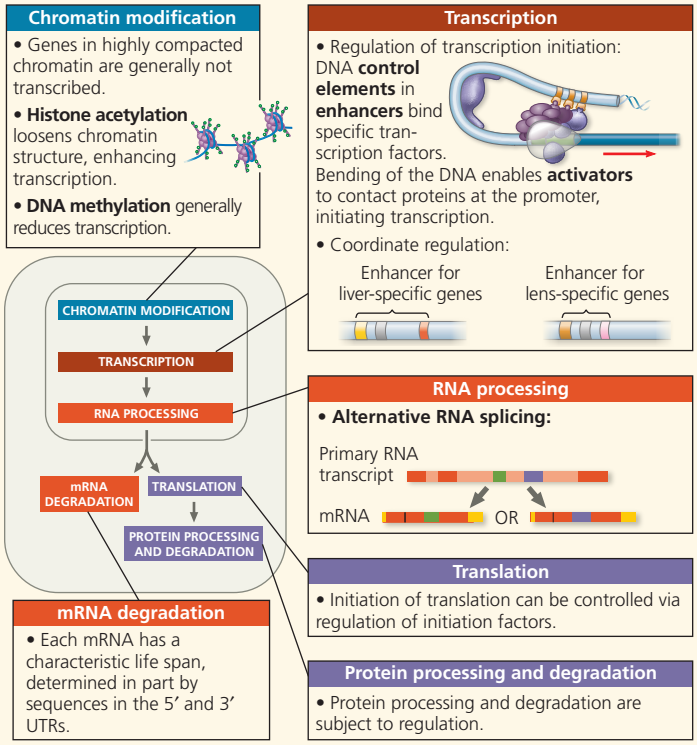


- Some operons are also subject to positive gene regulation via a stimulatory **activator** protein, such as cAMP receptor protein (CRP), which, when activated by **cyclic AMP**, binds to a site within the promoter and stimulates transcription.

? *Compare and contrast the roles of a corepressor and an inducer in negative regulation of an operon.*

CONCEPT 18.2

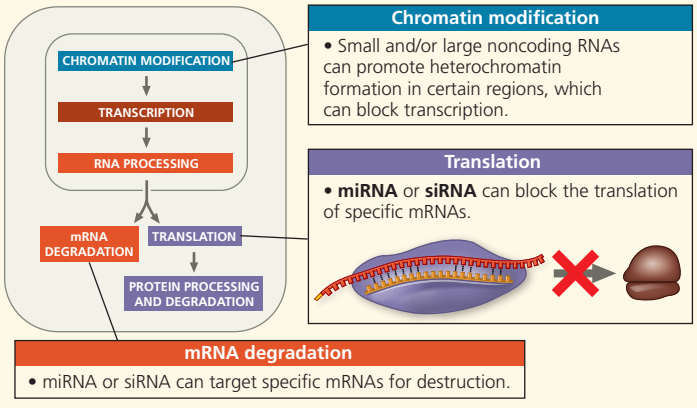
Eukaryotic gene expression is regulated at many stages *(pp. 418–426)*



? *Describe what must happen in a cell for a gene specific to that cell type to be transcribed.*

CONCEPT 18.3

Noncoding RNAs play multiple roles in controlling gene expression *(pp. 427–429)*



? *Why are miRNAs called noncoding RNAs? Explain how they participate in gene regulation.*

CONCEPT 18.4

A program of differential gene expression leads to the different cell types in a multicellular organism *(pp. 429–436)*

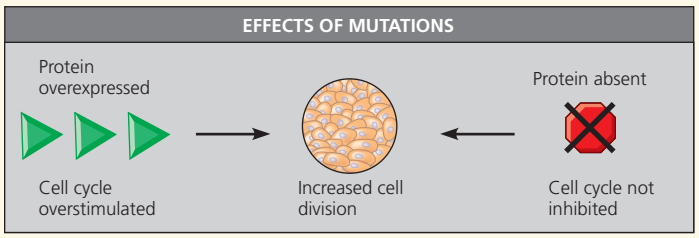
- Embryonic cells become committed to a certain fate (**determination**), and undergo **differentiation**, becoming specialized in structure and function for their determined fate. Cells differ in structure and function not because they contain different genomes but because they express different genes. **Morphogenesis** encompasses the processes that give shape to the organism and its various structures.
- Localized **cytoplasmic determinants** in the unfertilized egg are distributed differentially to daughter cells, where they regulate the expression of those cells' developmental fate. In the process called **induction**, signaling molecules from embryonic cells cause transcriptional changes in nearby target cells.
- Differentiation is heralded by the appearance of tissue-specific proteins, which enable differentiated cells to carry out their specialized roles.
- In animals, **pattern formation**, the development of a spatial organization of tissues and organs, begins in the early embryo. **Positional information**, the molecular cues that control pattern formation, tells a cell its location relative to the body's axes and to other cells. In *Drosophila*, gradients of **morphogens** encoded by **maternal effect genes** determine the body axes. For example, the gradient of **Bicoid** protein determines the anterior-posterior axis.

? *Describe the two main processes that cause embryonic cells to head down different pathways to their final fates.*

CONCEPT 18.5

Cancer results from genetic changes that affect cell cycle control *(pp. 436–442)*

- The products of **proto-oncogenes** and **tumor-suppressor genes** control cell division. A DNA change that makes a proto-oncogene excessively active converts it to an **oncogene**, which may promote excessive cell division and cancer. A tumor-suppressor gene encodes a protein that inhibits abnormal cell division. A mutation that reduces the activity of its protein product may lead to excessive cell division and cancer.
- Many proto-oncogenes and tumor-suppressor genes encode components of growth-stimulating and growth-inhibiting signaling pathways, respectively, and mutations in these genes can interfere with normal cell-signaling pathways. A hyperactive version of a protein in a stimulatory pathway, such as **Ras** (a G protein), functions as an oncogene protein. A defective version of a protein in an inhibitory pathway, such as **p53** (a transcription activator), fails to function as a tumor suppressor.



- In the multistep model of cancer development, normal cells are converted to cancer cells by the accumulation of mutations affecting proto-oncogenes and tumor-suppressor genes. Technical advances in DNA and mRNA sequencing are enabling cancer treatments that are more individually based.
- Genomics-based studies have resulted in researchers proposing four subtypes of breast cancer, based on expression of genes by tumor cells.
- Individuals who inherit a mutant allele of a proto-oncogene or tumor-suppressor gene have a predisposition to develop a particular cancer. Certain viruses promote cancer by integration of viral DNA into a cell's genome.

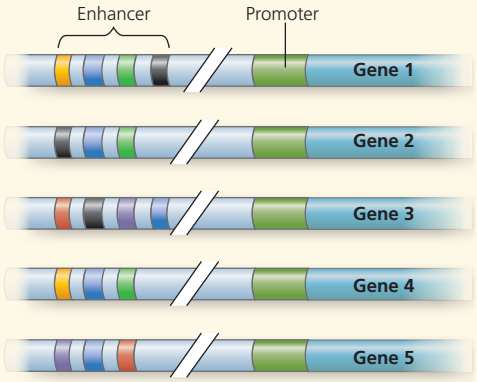
? *Compare the usual functions of proteins encoded by proto-oncogenes with those of proteins encoded by tumor-suppressor genes.*

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–10 can be found in the Study Area in MasteringBiology.

11. DRAW IT The diagram below shows five genes, including their enhancers, from the genome of a certain species. Imagine that yellow, blue, green, black, red, and purple activator proteins exist that can bind to the appropriately color-coded control elements in the enhancers of these genes.

PRACTICE TEST
goo.gl/iAsVgL



(a) Draw an X above enhancer elements (of all the genes) that would have activators bound in a cell in which only gene 5 is transcribed. Identify which colored activators would be present.
(b) Draw a dot above all enhancer elements that would have activators bound in a cell in which the green, blue, and yellow activators are present. Identify which gene(s) would be transcribed.
(c) Imagine that genes 1, 2, and 4 code for nerve-specific proteins, and genes 3 and 5 are skin specific. Identify which activators would have to be present in each cell type to ensure transcription of the appropriate genes.

12. EVOLUTION CONNECTION RNAi mechanisms, particularly the expression of miRNAs, are widely found in multicellular eukaryotes, but are not found in several unicellular eukaryotes, such as *Saccharomyces cerevisiae*. On the other hand, the mutation of Dicer, a protein essential for the generation of miRNA, in mice can be lethal for embryos. How can these concepts be reconciled?

13. SCIENTIFIC INQUIRY Prostate cells usually require testosterone and other androgens to survive. But some prostate cancer cells thrive despite treatments that eliminate androgens. One hypothesis is that estrogen, often considered a female hormone, may be activating genes normally controlled by an androgen in these cancer cells. Describe one or more experiments to test this hypothesis. (See Figure 9.9 to review the action of these steroid hormones.)

14. **SCIENCE, TECHNOLOGY, AND SOCIETY** Trace amounts of dioxin were present in Agent Orange, a defoliant sprayed on vegetation during the Vietnam War. Animal tests suggest that dioxin can cause birth defects, cancer, liver and thymus damage, and immune system suppression, sometimes leading to death. But the animal tests are equivocal; a hamster is not affected by a dose that can kill a guinea pig. Dioxin acts like a steroid hormone, entering a cell and binding to a cytoplasmic receptor that then binds the cell's DNA.
    (a) Discuss how this mechanism might help explain the variety of dioxin's effects on different body systems and in different animals.
    (b) Discuss how you might determine whether a type of illness is related to dioxin exposure. Next, discuss how you might determine whether a particular individual became ill as a result of exposure to dioxin. Which would be more difficult to demonstrate? Why?

15. **WRITE ABOUT A THEME: INTERACTIONS** In a short essay (100–150 words), discuss how the processes shown in Figure 18.2 are examples of feedback mechanisms regulating biological systems in bacterial cells.

16. **SYNTHESIZE YOUR KNOWLEDGE**

The flashlight fish has an organ under its eye that emits light, which serves to startle predators and attract prey, and allows the fish to communicate with other fish. Some species can rotate the organ inside and then out, so the light appears to flash on and off. The light is actually emitted by bacteria (of the genus *Vibrio*) that live in the organ in a mutualistic relationship with the fish. (The bacteria receive nutrients from the fish.) The bacteria must multiply until they reach a certain density in the organ (a "quorum"; see Concept 9.1), at which point they all begin emitting light at the same time. There is a group of six or so genes, called *lux* genes, whose gene products are necessary for light formation. Given that these bacterial genes are regulated together, propose a hypothesis for how the genes are organized and regulated.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

This page is intentionally left blank.

19 DNA Technology

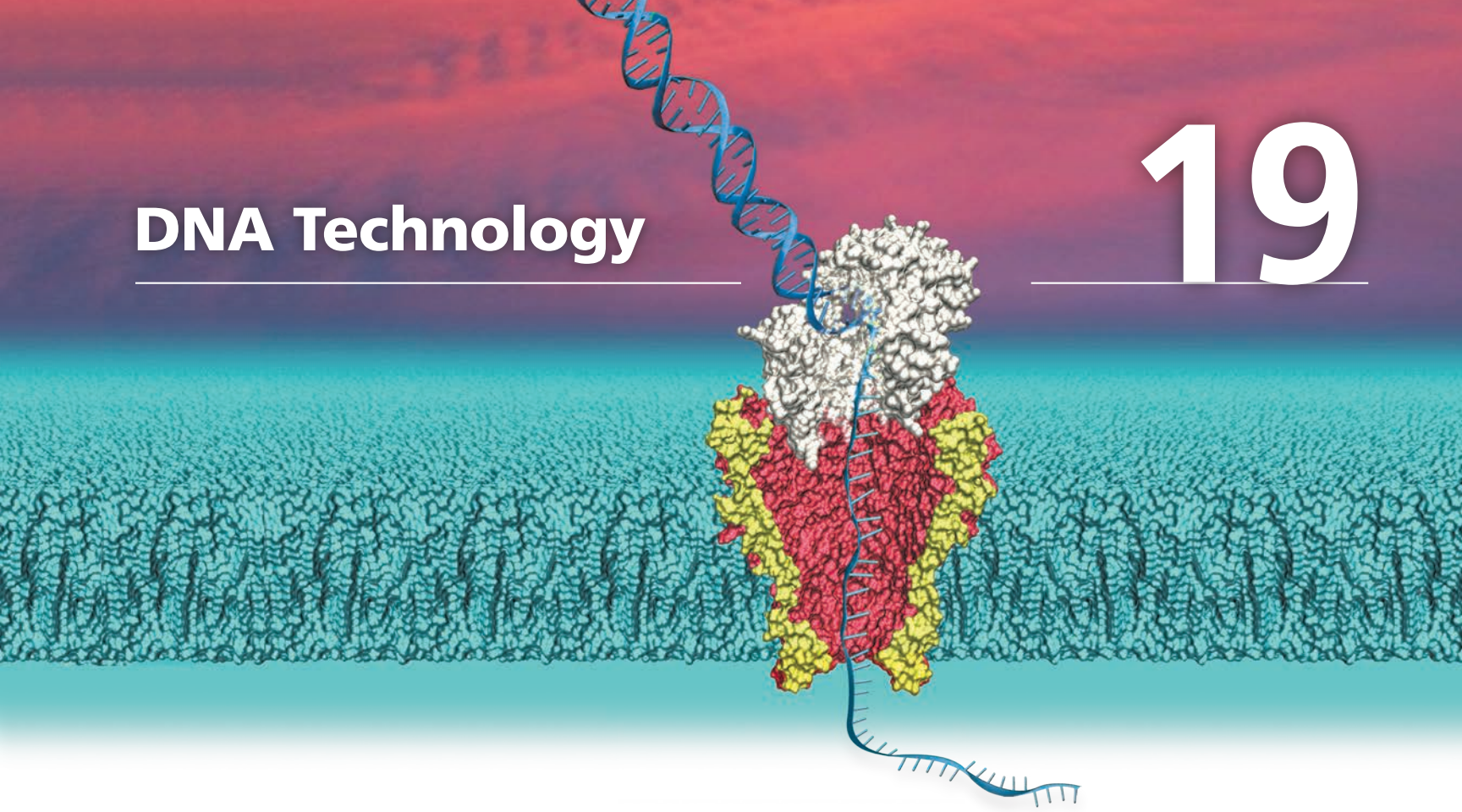


Figure 19.1 How can the technique shown in this model advance our sequencing of genomes?

KEY CONCEPTS

Woolly mammoth, an extinct organism whose genome was sequenced using mummified remains

The DNA Toolbox

The last decade or so has seen some extraordinary feats in biology, among them determination of the complete DNA sequences of several extinct species, including woolly mammoths (see below), Neanderthals, and a 700,000-year-old horse. Pivotal to those discoveries was the sequencing of the human genome, essentially completed in 2003. This endeavor marked a turning point in biology because it sparked remarkable technological advances in DNA sequencing.

The first human genome sequence took several years at a cost of 1 billion dollars; the time and cost of sequencing a genome have been in free fall since then. **Figure 19.1** shows a model of a sequencing technique in which the nucleotides of a single strand of DNA are passed one by one through a very small pore in a membrane, and the resulting tiny changes in an electrical current are used to determine the nucleotide sequence. Developers of this technique, which you will learn more about later in the chapter, claim that ultimately we will be able to sequence a human genome in about 6 hours on a $900 device the size of a pack of gum.

In this chapter, we'll first describe the main techniques for sequencing and manipulating DNA—**DNA technology**—and for using these DNA tools to analyze gene expression. Next, we'll explore advances in cloning organisms and producing stem cells, techniques that have both expanded our basic understanding of biology

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

and enhanced our ability to apply that understanding to global problems. In the last section, we'll survey the practical applications of DNA-based **biotechnology**, the manipulation of organisms or their components to make useful products. Today, the applications of DNA technology affect everything from agriculture to criminal law to medical research. We will end by considering some of the important social and ethical issues that arise as biotechnology becomes more pervasive in our lives.

CONCEPT 19.1

DNA sequencing and DNA cloning are valuable tools for genetic engineering and biological inquiry

The discovery of the structure of the DNA molecule, and specifically the recognition that its two strands are complementary to each other, opened the door for the development of DNA sequencing and other techniques used in biological research today. Key to these techniques is **nucleic acid hybridization**, the base pairing of one strand of a nucleic acid to a complementary sequence on a strand from a different nucleic acid molecule. In this section, we'll first describe DNA-sequencing techniques. Then we'll explore other important methods used in **genetic engineering**, the direct manipulation of genes for practical purposes.

DNA Sequencing

Researchers can exploit the principle of complementary base pairing to determine the complete nucleotide sequence of a DNA molecule, a process called **DNA sequencing**. The DNA is first cut into fragments, and then each fragment is sequenced. The first automated procedure used a technique called *dideoxyribonucleotide* (or *dideoxy*) *chain termination sequencing*. In this technique, one strand of a DNA fragment is used as a template for synthesis of a nested set of complementary fragments; these are further analyzed to yield the sequence. Biochemist Frederick Sanger received the Nobel Prize in 1980 for developing this method. Dideoxy sequencing is still used for routine small-scale sequencing jobs.

HHMI Video: Sanger Method of DNA Sequencing

In the last 15 years, "next-generation sequencing" techniques have been developed that are much faster **(Figure 19.2)**. DNA fragments are amplified (copied) to yield an enormous number of identical fragments **(Figure 19.3)**. A specific strand of each fragment is immobilized, and the complementary strand is synthesized, one nucleotide at a time. A chemical technique enables electronic monitors to identify in real time which of the four nucleotides is added; this method is thus called *sequencing by synthesis*. Thousands or hundreds of thousands of fragments, each about 300 nucleotides long, are sequenced in parallel in machines like those shown in Figure 19.2, accounting for the high rate of nucleotides sequenced per hour. This is an example of "high-throughput" DNA technology and is currently the method of choice for studies where massive numbers of DNA samples—even a set of numerous fragments representing an entire genome—are being sequenced.

Figure 19.2 Next-generation DNA sequencing machines. These machines use sequencing by synthesis and can sequence 70–90 million nucleotides per hour.

More and more often, next-generation sequencing is complemented (or in some cases replaced) by "third-generation sequencing," with each new technique being faster and less expensive than the previous one. In some new methods, the DNA is neither cut into fragments nor amplified. Instead, a single, very long DNA molecule is sequenced on its own. Several groups have developed techniques that move a single strand of a DNA molecule through a very small pore (a *nanopore*) in a membrane, identifying the bases one by one by the distinct way each interrupts an electrical current. One model of this concept is shown in Figure 19.1, in which the pore is a protein channel embedded in a lipid membrane. (Other researchers are using artificial membranes and nanopores.) The idea is that each type of base interrupts the electrical current for a slightly different length of time. In 2015, after a year of use and review by scientists, the first nanopore sequencer went on the market; this device is the size of a small candy bar and connects to a computer via a USB port. Associated software allows immediate identification and analysis of the sequence. This is only one of many approaches to further increase the rate and cut the cost of sequencing, while also allowing the methodology to move out of the laboratory and into the field.

Improved DNA-sequencing techniques have transformed the way in which we can explore fundamental biological questions about evolution and how life works (see Make Connections Figure 5.26). Little more than 15 years after the human genome sequence was announced, researchers had completed the sequencing of thousands of genomes, with tens of thousands in progress. Complete genome sequences have been determined for cells from several cancers, for ancient humans, and for the many bacteria that live in the

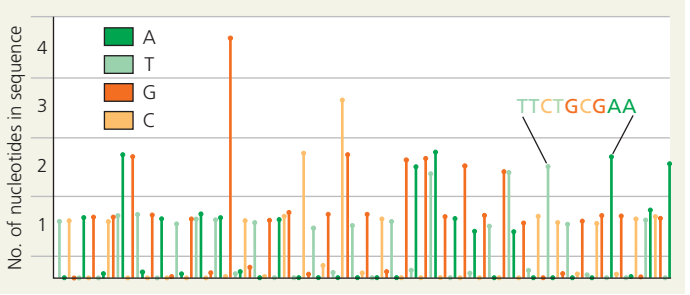
Figure 19.3

Research Method Sequencing by Synthesis: Next-Generation Sequencing

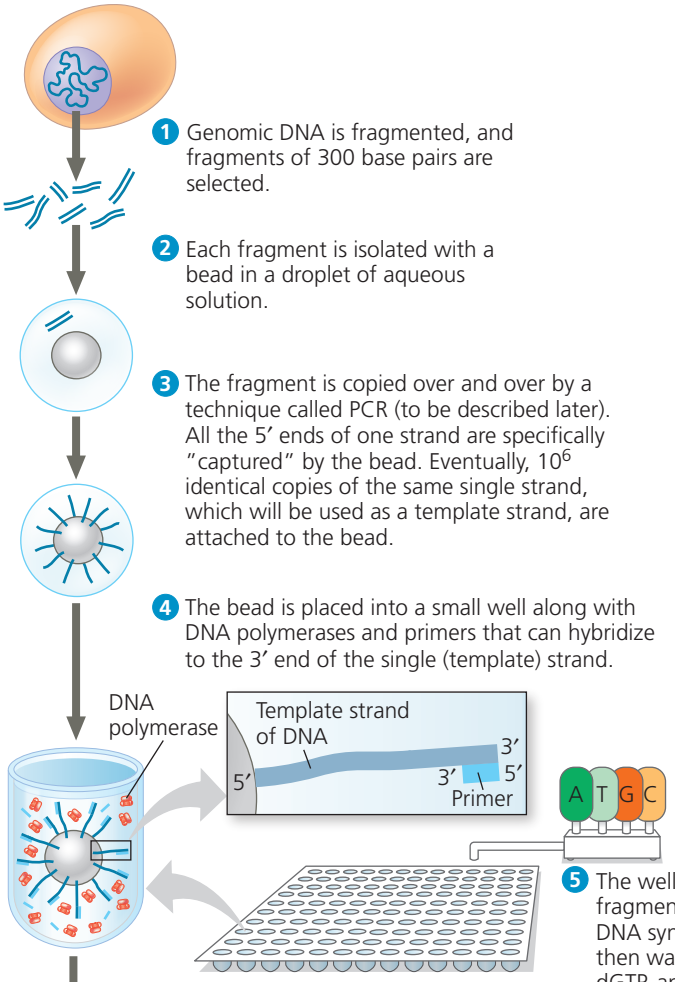
Application In current next-generation sequencing techniques, each fragment is about 300 nucleotides long; by sequencing the fragments in parallel, about 2 billion nucleotides can be sequenced in 24 hours.

Technique See numbered steps and diagrams.

Results Each of the 2,000,000 wells in the multiwell plate, which holds a different fragment, yields a different sequence. The results for one fragment are shown below as a "flow-gram." The sequences of the entire set of fragments are analyzed using computer software, which "stitches" them together into a whole sequence—here, an entire genome.



INTERPRET THE DATA ➤ *If the template strand has two or more identical nucleotides in a row, their complementary nucleotides will be added one after the other in the same flow step. How are two or more of the same nucleotide (in a row) detected in the flow-gram? (See sample on the right.) Write out the sequence of the first 25 nucleotides in the flow-gram above, starting from the left. (Ignore the very short lines.)*



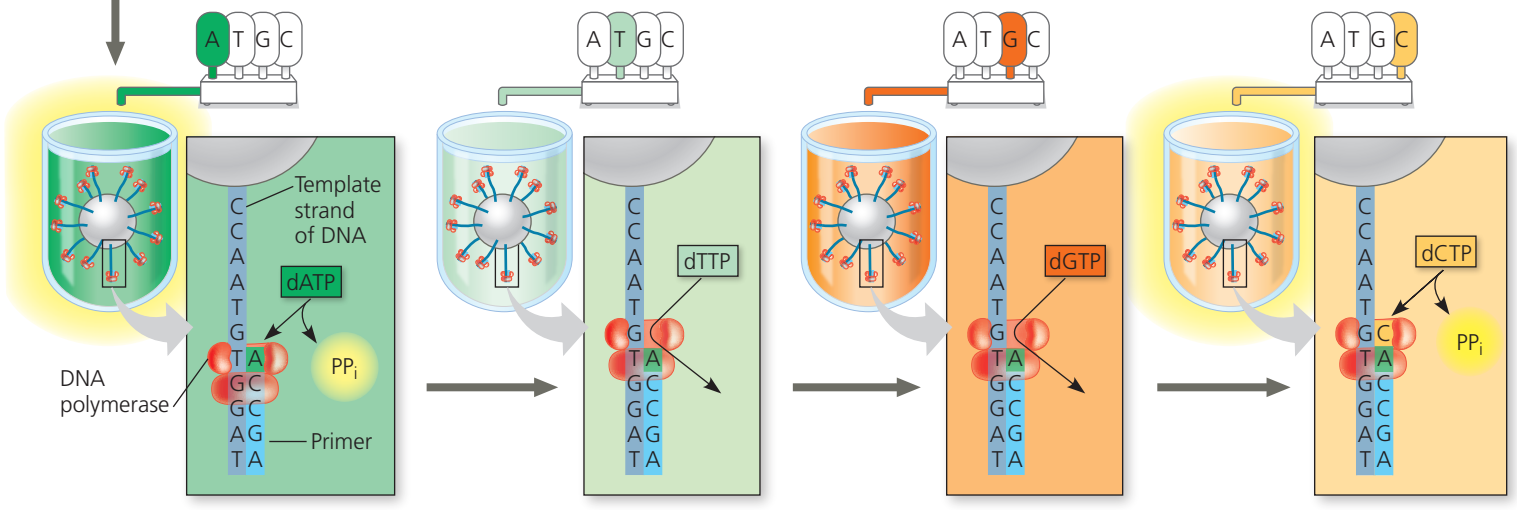
1 Genomic DNA is fragmented, and fragments of 300 base pairs are selected.

2 Each fragment is isolated with a bead in a droplet of aqueous solution.

3 The fragment is copied over and over by a technique called PCR (to be described later). All the 5′ ends of one strand are specifically "captured" by the bead. Eventually, 10^6 identical copies of the same single strand, which will be used as a template strand, are attached to the bead.

4 The bead is placed into a small well along with DNA polymerases and primers that can hybridize to the 3′ end of the single (template) strand.

5 The well is one of 2 million on a multiwell plate, each containing a different DNA fragment to be sequenced. A solution of one of the four nucleotides required for DNA synthesis (deoxynucleoside triphosphates, or dNTPs) is added to all wells and then washed off. This is done sequentially for all four nucleotides: dATP, dTTP, dGTP, and then dCTP. The entire process is then repeated.



6 In each well, if the next base on the template strand (T in this example) is complementary to the added nucleotide (A, here), the nucleotide is joined to the growing strand, releasing PP_i, which causes a flash of light that is recorded.

7 The nucleotide is washed off and a different nucleotide (dTTP, here) is added. If the nucleotide is not complementary to the next template base (G, here), it is not joined to the strand and there is no flash.

8 The process of adding and washing off the four nucleotides is repeated until every fragment has a complete complementary strand. The pattern of flashes reveals the sequence of the original fragment in each well.

human intestine. In Chapter 20, you'll learn more about how this rapid acceleration of sequencing technology has revolutionized our study of the evolution of species and the genome itself. Now, let's consider how individual genes are studied.

Making Multiple Copies of a Gene or Other DNA Segment

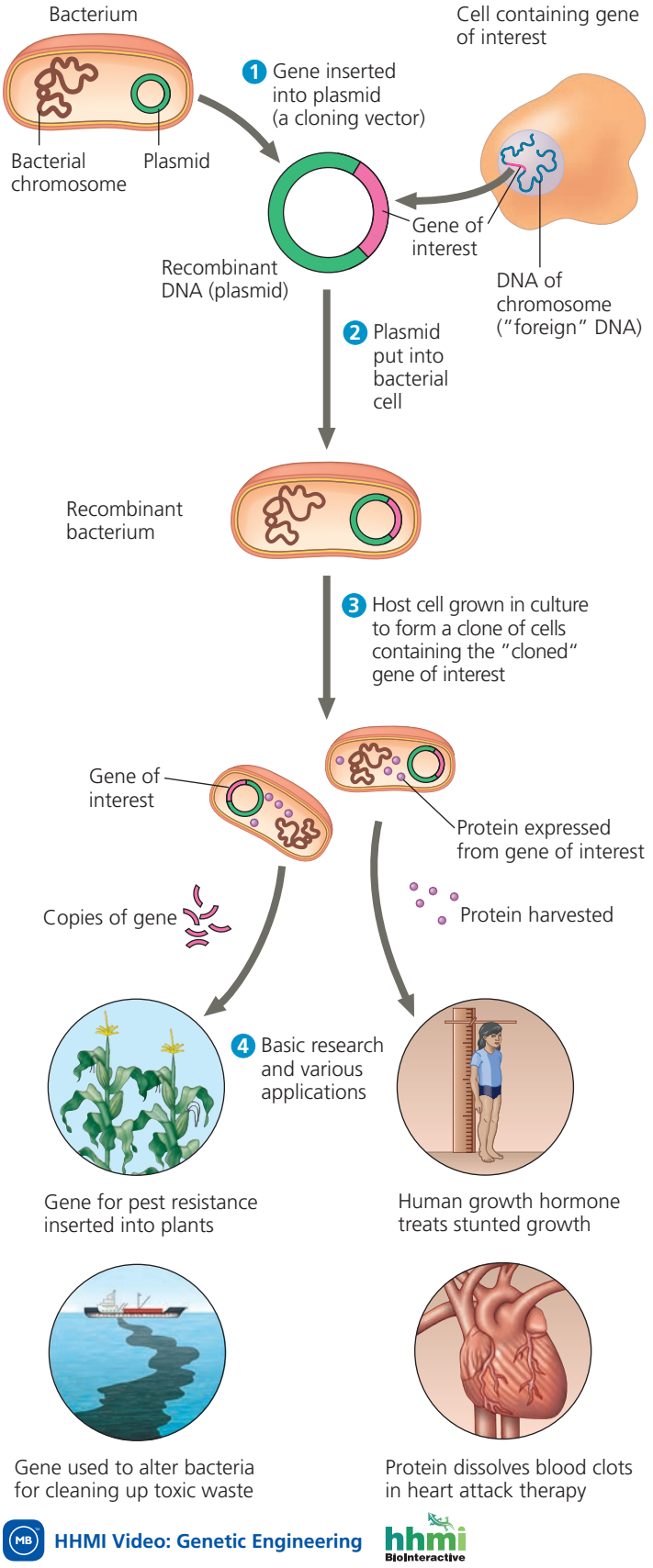
A molecular biologist studying a particular gene or group of genes faces a challenge. Naturally occurring DNA molecules are very long, and a single molecule usually carries hundreds or even thousands of genes. Moreover, in many eukaryotic genomes, protein-coding genes occupy only a small proportion of the chromosomal DNA, the rest being noncoding nucleotide sequences. A single human gene, for example, might constitute only 1/100,000 of a chromosomal DNA molecule. As a further complication, the distinctions between a gene and the surrounding DNA are subtle, consisting only of differences in nucleotide sequence. To work directly with specific genes, scientists have developed methods for preparing well-defined segments of DNA in multiple identical copies, a process called **DNA cloning**.

Most methods for cloning pieces of DNA in the laboratory share certain general features. One common approach uses bacteria, most often *Escherichia coli*. Recall from Figure 16.12 that the *E. coli* chromosome is a large, circular molecule of DNA. In addition, *E. coli* and many other bacteria also have **plasmids**, small, circular DNA molecules that are replicated separately. A plasmid has only a small number of genes; these genes may be useful when the bacterium is in a particular environment but may not be required for survival or reproduction under most conditions.

To clone pieces of DNA using bacteria, researchers first obtain a plasmid (originally isolated from a bacterial cell and genetically engineered for efficient cloning) and insert DNA from another source ("foreign" DNA) into it **(Figure 19.4)**. The resulting plasmid is now a **recombinant DNA molecule**, a molecule containing DNA from two different sources, very often different species. The plasmid is then returned to a bacterial cell, producing a *recombinant bacterium*. This single cell reproduces through repeated cell divisions to form a clone of cells, a population of genetically identical cells. Because the dividing bacteria replicate the recombinant plasmid and pass it on to their descendants, the foreign DNA and any genes it carries are cloned at the same time. The production of multiple copies of a single gene is a type of DNA cloning called **gene cloning**.

In Figure 19.4, the plasmid acts as a **cloning vector**, a DNA molecule that can carry foreign DNA into a host cell and be replicated there. Bacterial plasmids are widely used as cloning vectors for several reasons: They can be readily obtained from commercial suppliers, manipulated to form recombinant plasmids by insertion of foreign DNA in a test

Figure 19.4 Gene cloning and some uses of cloned genes. In this simplified diagram of gene cloning, we start with a plasmid (originally isolated from a bacterial cell) and a gene of interest from another organism. Only one plasmid and one copy of the gene of interest are shown at the top of the figure, but the starting materials would include many of each.



HHMI Video: Genetic Engineering

tube (referred to as *in vitro*, from the Latin meaning "in glass"), and then easily introduced into bacterial cells. The foreign DNA in Figure 19.4 is a gene from a eukaryotic cell; we will describe in more detail how the foreign DNA segment was obtained later in this section.

Gene cloning is useful for two basic purposes: to make many copies of, or *amplify*, a particular gene and to produce a protein product from it (see Figure 19.4). Researchers can isolate copies of a cloned gene from bacteria for use in basic research or to endow another organism with a new metabolic capability, such as pest resistance. For example, a resistance gene present in one crop species might be cloned and transferred into plants of another species. (Such organisms are called *genetically modified,* or *GM* for short; they will be discussed later in the chapter.) Alternatively, a protein with medical uses, such as human growth hormone, can be harvested in large quantities from cultures of bacteria carrying a cloned gene for the protein. (We'll describe the techniques for expressing cloned genes later.) Since one gene is only a very small part of the total DNA in a cell, the ability to amplify such rare DNA fragments is crucial for any application involving a single gene.

Using Restriction Enzymes to Make a Recombinant DNA Plasmid

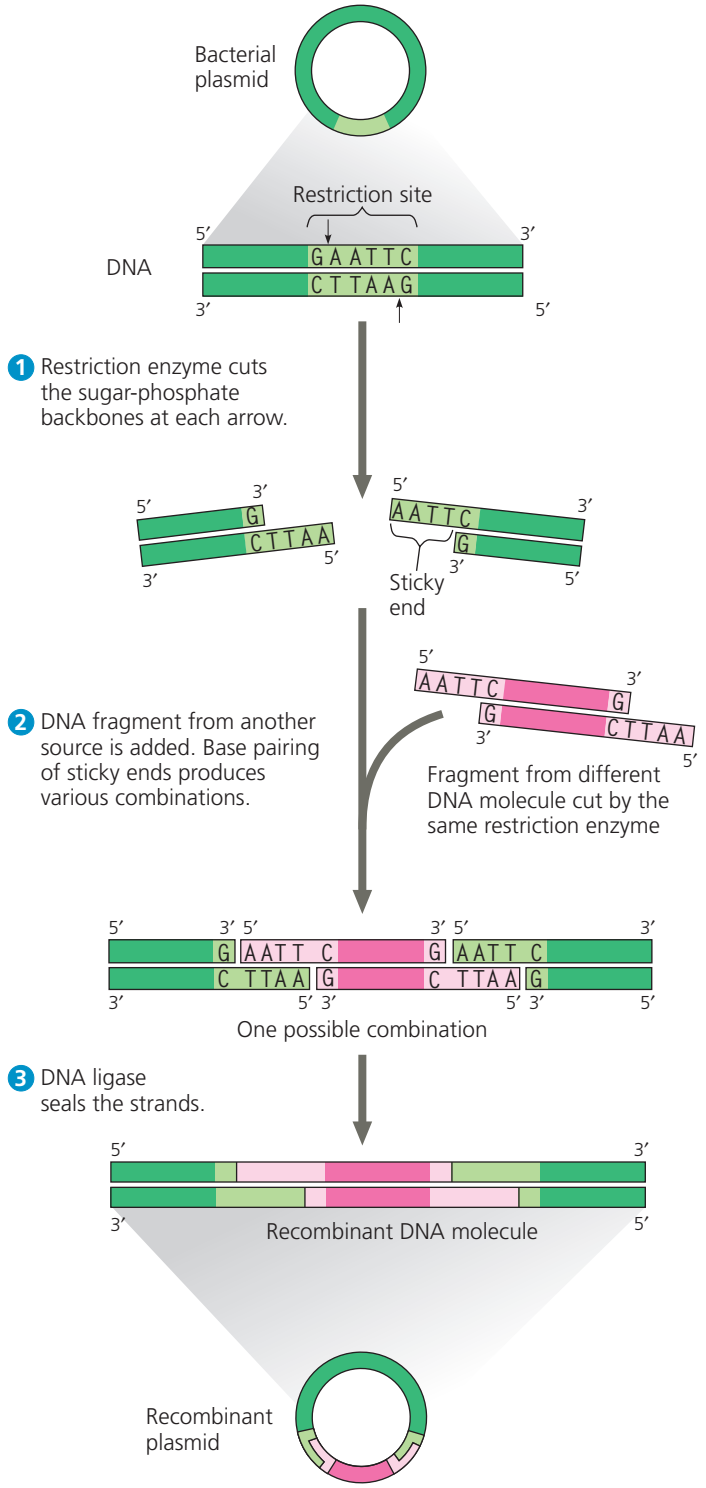
Gene cloning and genetic engineering generally rely on the use of enzymes that cut DNA molecules at a limited number of specific locations. These enzymes, called restriction endonucleases, or **restriction enzymes**, were discovered in the late 1960s by biologists doing basic research on bacteria. Restriction enzymes protect the bacterial cell by cutting up foreign DNA from other organisms or phages (see Concept 26.2).

Hundreds of different restriction enzymes have been identified and isolated. Each restriction enzyme is very specific, recognizing a particular short DNA sequence, or **restriction site**, and cutting both DNA strands at precise points within this restriction site. The DNA of a bacterial cell is protected from the cell's own restriction enzymes by the addition of methyl groups ($-CH_3$) to adenines or cytosines within the sequences recognized by the enzymes.

Animation: Restriction Enzymes

Figure 19.5 shows how restriction enzymes are used to clone a foreign DNA fragment into a bacterial plasmid. At the top is a bacterial plasmid (like the one in Figure 19.4) that has a single restriction site recognized by a particular restriction enzyme from *E. coli*. As shown in this example, most restriction sites are symmetrical. That is, the sequence of nucleotides is the same on both strands when read in the 5′ → 3′ direction. The most commonly used restriction enzymes recognize sequences containing four to eight nucleotide pairs. Because any sequence this short usually occurs (by chance)

Figure 19.5 Using a restriction enzyme and DNA ligase to make a recombinant DNA plasmid. The restriction enzyme in this example (called *Eco*RI) recognizes a single six-base-pair restriction site present in this plasmid. It makes staggered cuts in the sugar-phosphate backbones, producing fragments with "sticky ends." Foreign DNA fragments with complementary sticky ends can base-pair with the plasmid ends; the ligated product is a recombinant plasmid. (If the two plasmid sticky ends base-pair, the original nonrecombinant plasmid is reformed.)



DRAW IT ➤ *The restriction enzyme* Hind*III recognizes the sequence 5′-AAGCTT-3′, cutting between the two A's. Draw the double-stranded sequence before and after the enzyme cuts it.*

Animation: Recombinant DNA

many times in a long DNA molecule, a restriction enzyme will make many cuts in such a DNA molecule, yielding a set of **restriction fragments**. All copies of a given DNA molecule always yield the same set of restriction fragments when exposed to the same restriction enzyme.

The most useful restriction enzymes cleave the sugar-phosphate backbones in the two DNA strands in a staggered manner, as indicated in Figure 19.5. The resulting double-stranded restriction fragments have at least one single-stranded end, called a **sticky end**. These short extensions can form hydrogen-bonded base pairs with complementary sticky ends on any other DNA molecules cut with the same enzyme. The associations formed in this way are only temporary but can be made permanent by DNA ligase, which catalyzes the formation of covalent bonds that close up the sugar-phosphate backbones of DNA strands (see Figure 16.16). You can see at the bottom of Figure 19.5 that the ligase-catalyzed joining of DNA from two different sources produces a stable recombinant DNA molecule, in this example, a recombinant plasmid.

MB **Animation: Creating Recombinant DNA**

To check the recombinant plasmids after they have been copied many times in host cells (see Figure 19.4), a researcher might cut the products again using the same restriction enzyme, expecting two DNA fragments, one the size of the plasmid and one the size of the inserted DNA. To separate and visualize the fragments, researchers carry out a technique called **gel electrophoresis**, which uses a gel made of a polymer as a molecular sieve to separate out a mixture of nucleic acid fragments by length **(Figure 19.6)**. Gel electrophoresis is used in conjunction with many different techniques in molecular biology.

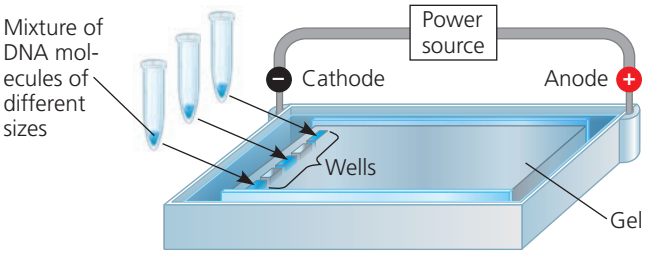
Now that we have discussed the cloning vector in some detail, let's consider the foreign DNA to be inserted. The most common way to obtain many copies of the gene to be cloned is by PCR, described next.

Amplifying DNA: The Polymerase Chain Reaction (PCR) and Its Use in DNA Cloning

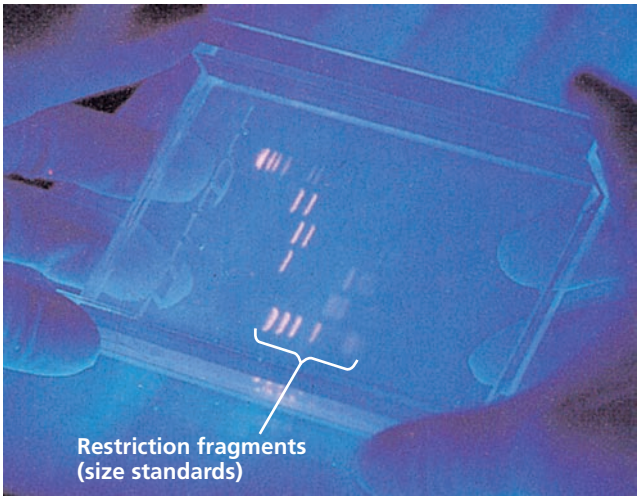
Today, most researchers have some information about the sequence of the gene or other DNA segment they want to clone. Using this information, they can start with the entire collection of genomic DNA from the particular species of interest and obtain many copies of the desired gene by using a technique called the **polymerase chain reaction**, or **PCR**. **Figure 19.7** illustrates the steps in PCR. Within a few hours, this technique can make billions of copies of a specific target DNA segment in a sample, even if that segment makes up less than 0.001% of the total DNA in the sample.

MB **HHMI Video: Polymerase Chain Reaction (PCR)** hhmi BioInteractive

Figure 19.6 Gel electrophoresis. A gel made of a polymer acts as a molecular sieve to separate nucleic acids or proteins differing in size, electrical charge, or other physical properties as they move in an electric field. In the example shown here, DNA molecules are separated by length in a gel made of a polysaccharide called agarose.



(a) Each sample, a mixture of different DNA molecules, is placed in a separate well near one end of a thin slab of agarose gel. The gel is set into a small plastic support and immersed in an aqueous, buffered solution in a tray with electrodes at each end. The current is then turned on, causing the negatively charged DNA molecules to move toward the positive electrode.



(b) Shorter molecules are slowed down less than longer ones, so they move faster through the gel. After the current is turned off, a DNA-binding dye is added that fluoresces pink in UV light. Each pink band corresponds to many thousands of DNA molecules of the same length. The horizontal ladder of bands at the bottom of the gel is a set of restriction fragments of known sizes for comparison with samples of unknown length.

MB **Animation: Gel Electrophoresis of DNA**

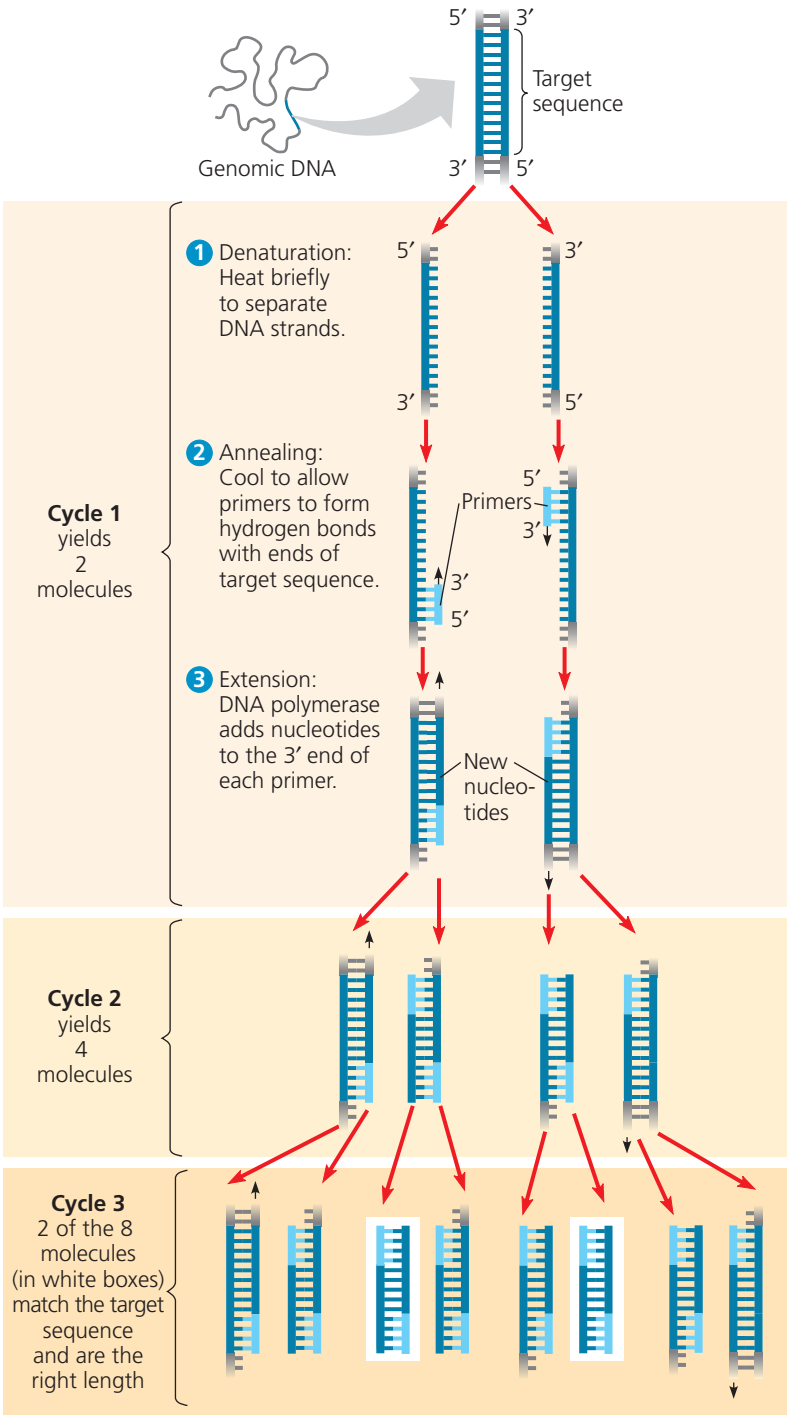
In the PCR procedure, a three-step cycle brings about a chain reaction that produces an exponentially growing population of identical DNA molecules. During each cycle, the reaction mixture is heated to denature (separate) the strands of the double-stranded DNA and then cooled to allow annealing (hydrogen bonding) of short, single-stranded DNA primers complementary to sequences on opposite strands at each end of the target sequence; finally, a heat-stable DNA polymerase extends the primers in the 5′ → 3′ direction. If a standard DNA polymerase were used, the protein would be denatured along with the DNA during the first heating step and would have to be replaced after each cycle. The key to

Figure 19.7

Research Method The Polymerase Chain Reaction (PCR)

Application With PCR, any specific segment—the so-called target sequence—in a DNA sample can be copied many times (amplified) within a test tube.

Technique PCR requires double-stranded DNA containing the target sequence, a heat-resistant DNA polymerase, all four nucleotides, and two 15- to 20-nucleotide single DNA strands that serve as primers. One primer is complementary to one end of the target sequence on one strand; the second primer is complementary to the other end of the sequence on the other strand.



Results After three cycles, two molecules match the target sequence exactly. After 30 more cycles, over 1 billion (10^9) molecules match the target sequence.

Animation: Copying DNA through PCR

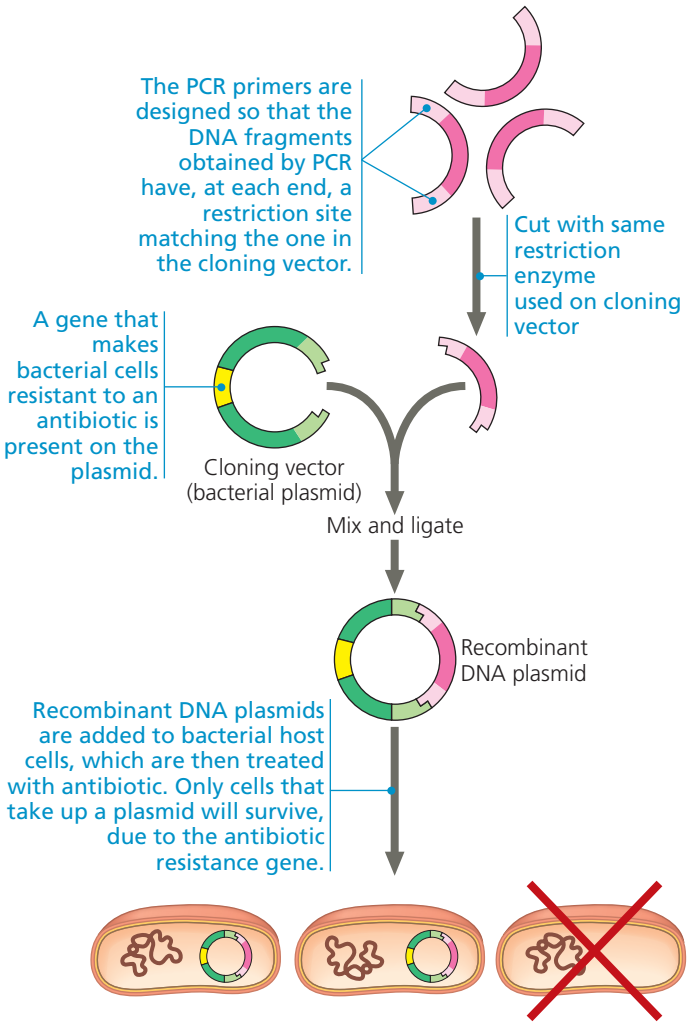
automating PCR was the discovery of an unusual heat-stable DNA polymerase called *Taq* polymerase, named after the bacterial species from which it was first isolated. This bacterial species, *Thermus aquaticus*, lives in hot springs, and the stability of its DNA polymerase at high temperatures is an evolutionary adaptation that enables the enzyme to function at temperatures up to 95°C. Today, researchers also use a DNA polymerase from the archaean species *Pyrococcus furiosus*. This enzyme, called *Pfu* polymerase, is more accurate and stable but more expensive than *Taq* polymerase.

PCR is speedy and very specific. Only a minuscule amount of DNA need be present in the starting material, and this DNA can be partially degraded, as long as there are a few copies of the complete target sequence. The key to the high specificity is the pair of primers used for each PCR amplification. The primer sequences are chosen so they hybridize *only* to sequences at opposite ends of the target segment, one on the 3′ end of each strand. (For high specificity, the primers must be at least 15 or so nucleotides long.) With each successive cycle, the number of target segment molecules of the correct length doubles, so the number of molecules equals 2^n, where n is the number of cycles. After 30 or so cycles, about a billion copies of the target sequence are present!

Despite its speed and specificity, PCR amplification cannot substitute for gene cloning in cells to make large amounts of a gene. This is because occasional errors during PCR replication limit the number of good copies and the length of DNA fragments that can be copied. Instead, PCR is used to provide the specific DNA fragment for cloning. PCR primers are synthesized to include a restriction site at each end of the DNA fragment that matches the site in the cloning vector, and the fragment and vector are cut and ligated together **(Figure 19.8)**. The resulting plasmids are sequenced so that those with error-free inserts can be selected.

Devised in 1985, PCR has had a major impact on biological research and genetic engineering. PCR has been used to amplify DNA from a wide variety of sources: a 40,000-year-old frozen woolly mammoth (see the photo on the first page of this chapter); fingerprints or tiny amounts of blood, tissue, or semen found at crime scenes; single embryonic cells for rapid prenatal diagnosis of genetic disorders (see Figure 14.19); and cells infected with viruses that are difficult to detect, such as HIV. (To test for HIV, viral genes are amplified.) We'll return to applications of PCR later.

Figure 19.8 Use of a restriction enzyme and PCR in gene cloning. In a closer look at the process shown at the top of Figure 19.4, PCR is used to produce the DNA fragment or gene of interest that will be ligated into a cloning vector, in this case a bacterial plasmid. The ends of the fragments have the same restriction site as the cloning vector. The plasmid and the DNA fragments are cut with the same restriction enzyme and combined so the sticky ends can hybridize and be ligated together. The resulting plasmids are then introduced into bacterial host cells. The plasmid also contains an antibiotic resistance gene that allows only cells with a plasmid to survive when the antibiotic is present. Other genetic engineering techniques are used to ensure that cells with nonrecombinant plasmids can be eliminated.



Expressing Cloned Eukaryotic Genes

Once a gene has been cloned in host cells, its protein product can be expressed in large amounts for research or for practical applications, which we'll explore in Concept 19.4. Cloned genes can be expressed in either bacterial or eukaryotic cells; each option has advantages and disadvantages.

Bacterial Expression Systems

Getting a cloned eukaryotic gene to function in bacterial host cells can be difficult because certain aspects of gene expression are different in eukaryotes and bacteria. To overcome differences in promoters and other DNA control sequences, scientists usually employ an **expression vector**, a cloning vector that contains a highly active bacterial promoter just upstream of a restriction site where the eukaryotic gene can be inserted in the correct reading frame. The bacterial host cell will recognize the promoter and proceed to express the foreign gene now linked to that promoter. Such expression vectors allow the synthesis of many eukaryotic proteins in bacterial cells.

Another problem with expressing cloned eukaryotic genes in bacteria is the presence of noncoding regions (introns) in most eukaryotic genes (see Concept 17.3). Introns can make a eukaryotic gene very long and unwieldy, and they prevent correct expression of the gene by bacterial cells, which do not have RNA-splicing machinery. This problem can be surmounted by using a form of the gene that includes only the exons. (This is called *complementary DNA,* or *cDNA*; see Figure 19.10.)

Eukaryotic DNA Cloning and Expression Systems

Molecular biologists can avoid eukaryotic-bacterial incompatibility by using eukaryotic cells such as yeasts as hosts for cloning and expressing eukaryotic genes. Yeasts, single-celled fungi, are as easy to grow as bacteria, and they have plasmids, a rarity among eukaryotes.

In addition to enabling RNA splicing, eukaryotic host cells are advantageous because many eukaryotic proteins will not function unless they are modified after translation—for example, by the addition of carbohydrate groups (glycosylation) or lipid groups. Bacterial cells cannot carry out these modifications, and if the gene product requiring such processing is from a mammal, even yeast cells may not be able to modify the protein correctly. Several cultured cell types have proved successful as host cells for this purpose, including some mammalian cell lines and an insect cell line that can be infected by a virus (called baculovirus) carrying recombinant DNA.

Besides using vectors, scientists have developed other methods for introducing recombinant DNA into eukaryotic cells. In **electroporation**, a brief electrical pulse applied to a solution containing cells creates temporary holes in their plasma membranes, through which DNA can enter. (This technique is now commonly used for bacteria as well.) Alternatively, scientists can inject DNA directly into single eukaryotic cells using microscopically thin needles. Another way to get DNA into plant cells is by using the soil bacterium *Agrobacterium tumefaciens*, as we'll discuss later. Whatever the method, if the introduced DNA is incorporated into a cell's genome by genetic recombination, then it can be expressed by the cell. Expressing different versions of genes in cells allows researchers to study protein function, a topic we'll return to in Concept 19.2.

Cross-Species Gene Expression and Evolutionary Ancestry

EVOLUTION The ability to express eukaryotic proteins in bacteria (even if the proteins aren't modified properly) is quite remarkable when we consider how different eukaryotic and bacterial cells are. In fact, examples abound of genes that are taken from one species and function perfectly well when transferred into another very different species, such as a firefly gene in a tobacco plant and a jellyfish gene in a pig (see Figure 17.7). These observations underscore the shared evolutionary ancestry of species living today.

One example involves a gene called *Pax-6*, which has been found in animals as diverse as vertebrates and fruit flies. The vertebrate *Pax-6* gene product (the PAX-6 protein) triggers a complex program of gene expression resulting in formation of the vertebrate eye, which has a single lens. Expression of the fly *Pax-6* gene leads to formation of the compound fly eye, which is quite different from the vertebrate eye. When the mouse *Pax-6* gene was cloned and introduced into a fly embryo so that it replaced the fly's own *Pax-6* gene, researchers were surprised to see that the mouse version of the gene led to formation of a compound fly eye (see Figure 50.16). Conversely, when the fly *Pax-6* gene was transferred into a vertebrate embryo—a frog, in this case—a frog eye formed. Although the genetic programs triggered in vertebrates and flies generate very different eyes, the two versions of the *Pax-6* gene can substitute for each other to trigger lens development, evidence of their evolution from a gene in a very ancient common ancestor. Because of their ancient evolutionary roots, all living organisms share the same basic mechanisms of gene expression. This commonality is the basis of many recombinant DNA techniques described in this chapter.

CONCEPT CHECK 19.1

1. **MAKE CONNECTIONS ➤** The restriction site for an enzyme called *Pvu*I is the following sequence:

   5′-CGATCG-3′
   3′-GCTAGC-5′

   Staggered cuts are made between the T and C on each strand. What type of bonds are being cleaved? (See Concept 5.5.)
2. **DRAW IT ➤** One strand of a DNA molecule has the following sequence:

   5′-CTTGACGATCGTTACCG-3′

   Draw the other strand. Will *Pvu*I (see question 1) cut this molecule? If so, draw the products.
3. What are some potential difficulties in using plasmid vectors and bacterial host cells to produce large quantities of proteins from cloned eukaryotic genes?
4. **VISUAL SKILLS ➤** Compare Figure 19.7 with Figure 16.20. How does replication of DNA ends during PCR proceed without shortening the fragments each time?

For suggested answers, see Appendix A.

CONCEPT 19.2

Biologists use DNA technology to study gene expression and function

To see how a biological system works, scientists seek to understand the functioning of the system's component parts. Analysis of when and where a gene or group of genes is expressed can provide important clues about their function.

Analyzing Gene Expression

Biologists driven to understand the assorted cell types of a multicellular organism, cancer cells, or the developing tissues of an embryo first try to discover which genes are expressed by the cells of interest. The most straightforward way to do this is usually to identify the mRNAs being made. We'll first examine techniques that look for patterns of expression of specific individual genes. Next, we'll explore ways to characterize groups of genes being expressed by cells or tissues of interest. As you will see, all of these procedures depend in some way on base pairing between complementary nucleotide sequences.

Studying the Expression of Single Genes

Suppose we have cloned a gene that we suspect plays an important role in the embryonic development of *Drosophila melanogaster* (the fruit fly). The first thing we might want to know is which embryonic cells express the gene—in other words, where in the embryo is the corresponding mRNA found? We can detect the mRNA by nucleic acid hybridization with molecules of complementary sequence that we can follow in some way. The complementary molecule, a short, single-stranded nucleic acid that can be either RNA or DNA, is called a **nucleic acid probe**. Using our cloned gene as a template, we can synthesize a probe complementary to the mRNA. For example, if part of the sequence on the mRNA were

5′ ···C U C A U C A C C G G C··· 3′

then we would synthesize this single-stranded DNA probe:

3′ G A G T A G T G G C C G 5′

Each probe molecule is labeled during synthesis with a fluorescent tag so we can follow it. A solution containing probe molecules is applied to *Drosophila* embryos, allowing the probe to hybridize specifically with any complementary sequences on the many mRNAs in embryonic cells in which the gene is being transcribed. Because this technique allows us to see the mRNA in place (or *in situ*, in Latin) in the intact organism, this technique is called ***in situ* hybridization**. Different probes can be labeled with different fluorescent dyes, sometimes with strikingly beautiful results **(Figure 19.9)**.

Figure 19.9 Determining where single genes are expressed by *in situ* hybridization analysis. A *Drosophila* embryo was incubated in a solution containing DNA probes for five different mRNAs, each probe labeled with a different fluorescently colored tag. The embryo was then viewed using fluorescence microscopy; the resulting fluorescent micrograph is shown. Each color marks where a specific gene is expressed as mRNA. The arrows from the groups of yellow and blue cells above the micrograph show a magnified view of nucleic acid hybridization of the appropriately colored probe to the mRNA. Yellow cells (expressing the *wg* gene) interact with blue cells (expressing the *en* gene); their interaction helps establish the pattern in a body segment. The diagram at the bottom clarifies the eight segments visible in this view.

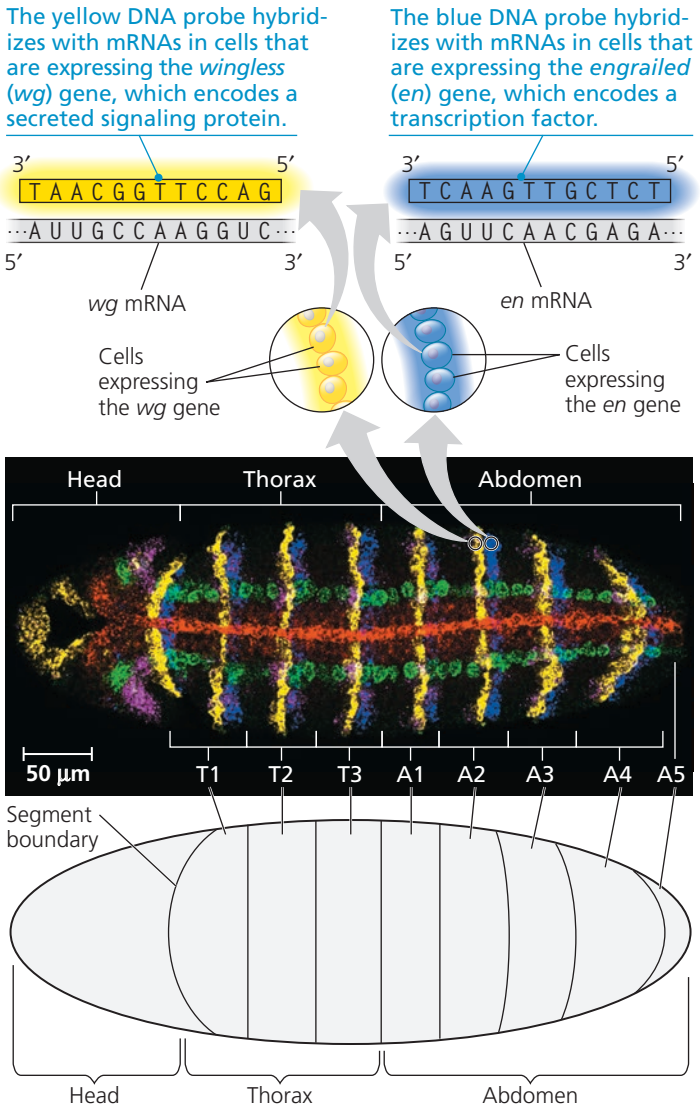
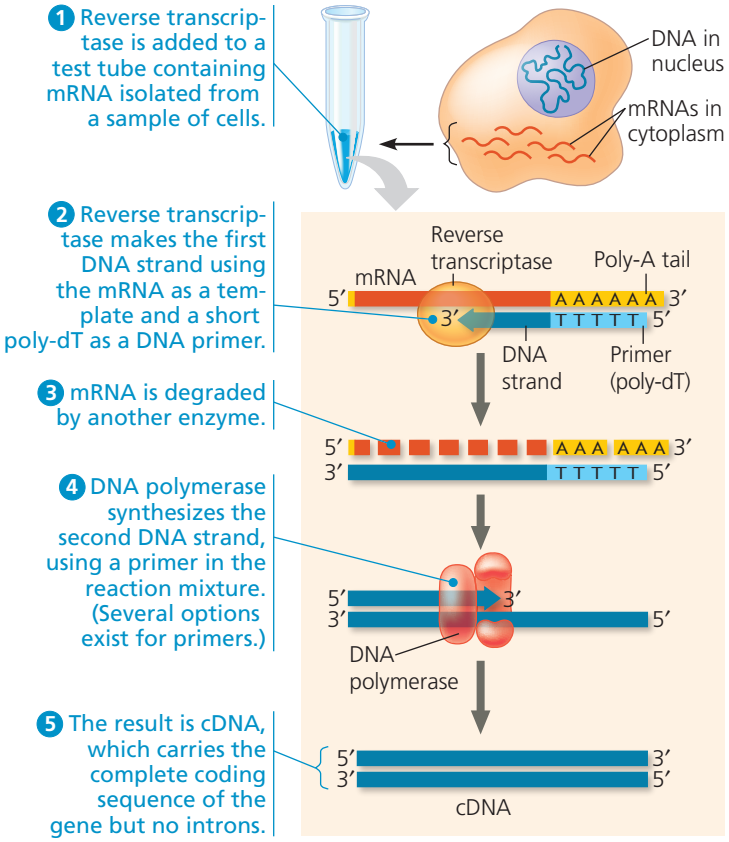


Figure 19.10 Making complementary DNA (cDNA) from eukaryotic genes. Complementary DNA is made in a test tube using mRNA as a template for the first strand. Only one mRNA is shown here, but the final collection of cDNAs would reflect all the mRNAs present in the cell.



Other mRNA detection techniques may be preferable for comparing the amounts of a specific mRNA in several samples at the same time—for example, in different cell types or in embryos at different stages of development. One method that is widely used is called the **reverse transcriptase polymerase chain reaction**, or **RT-PCR**.

RT-PCR begins by turning sample sets of mRNAs into double-stranded DNAs with the corresponding sequences. First, the enzyme reverse transcriptase (from a virus; see Figure 26.9) is used to synthesize a complementary DNA copy of each mRNA in the sample, called a *reverse transcript* **(Figure 19.10)**. Recall that the 3′ end of an mRNA has a stretch of adenine (A) nucleotides called a poly-A tail. This allows a short complementary strand of thymine deoxyribonucleotides (poly-dT) to be added and used as a primer for synthesis of this DNA strand. Following enzymatic degradation of the mRNA, a second DNA strand, complementary to the first, is synthesized by DNA polymerase. The resulting double-stranded DNA is called **complementary DNA (cDNA)**. (Made from mRNA, cDNA lacks introns and can be used for protein expression in bacteria, as mentioned earlier.) To analyze the timing of expression of the *Drosophila* gene of interest, for example, we would first isolate all the mRNAs from different stages of *Drosophila* embryos and make cDNA from each stage **(Figure 19.11)**.

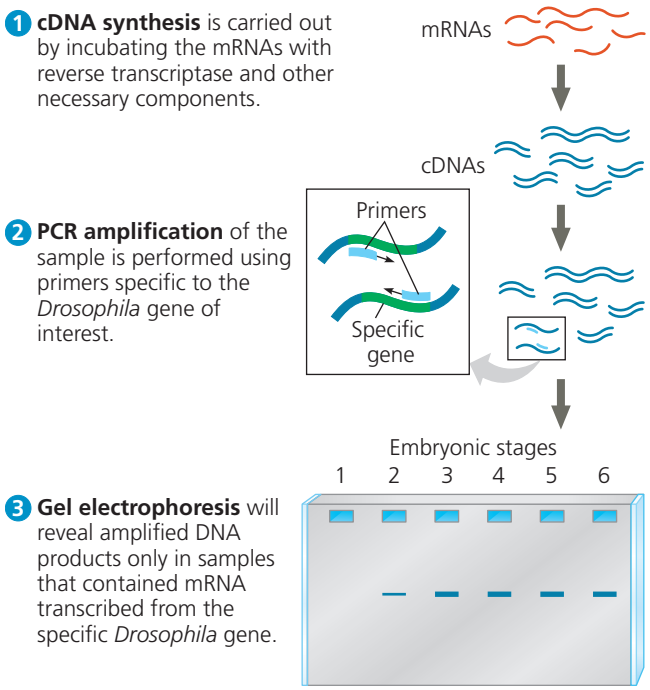
Next in RT-PCR is the PCR step (see Figure 19.7). As you will recall, PCR is a way of rapidly making many copies of one specific stretch of double-stranded DNA, using primers that hybridize to the opposite ends of the segment of interest. In our case, we would add primers corresponding to a segment of our *Drosophila* gene, using the cDNA from each embryonic stage as a template for PCR amplification in separate samples.

Figure 19.11

Research Method RT-PCR Analysis of the Expression of Single Genes

Application RT-PCR uses the enzyme reverse transcriptase (RT) in combination with PCR and gel electrophoresis. RT-PCR can be used to compare gene expression between samples—for instance, in different embryonic stages, in different tissues, or in the same type of cell under different conditions.

Technique In this example, samples containing mRNAs from six embryonic stages of *Drosophila* were analyzed for a specific mRNA as shown below. (In steps 1 and 2, the mRNA from only one stage is shown.)



Results The mRNA for this gene first is expressed at stage 2 and continues to be expressed through stage 6. The size of the amplified fragment (shown by its position on the gel) depends on the distance between the primers that were used (not on the size of the mRNA).

When the products are analyzed on a gel, copies of the amplified region will be observed as bands only in samples that originally contained mRNA from the gene of interest. An enhancement called *quantitative RT-PCR (qRT-PCR)* uses a fluorescent dye that fluoresces only when bound to a double-stranded PCR product. The newer quantitative PCR machines can detect the light and measure the PCR product, thus avoiding the need for electrophoresis while also providing quantitative data, a distinct advantage. RT-PCR or qRT-PCR can also be carried out with mRNAs collected from different tissues at one time to discover which tissue is producing a specific mRNA.

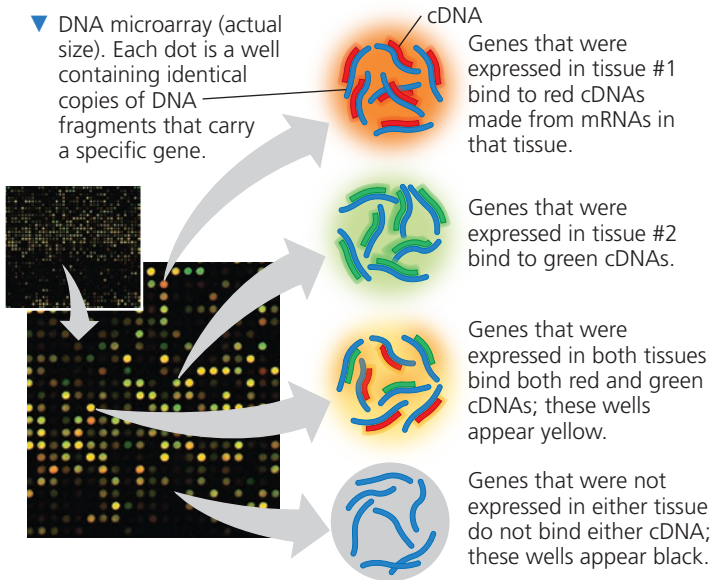
MB **Instructors**: The **Scientific Skills Exercise** "Analyzing Quantitative and Spatial Gene Expression Data" can be assigned in MasteringBiology. Students work with data from an experiment that investigated mRNA expression using both *in situ* hybridization and quantitative RT-PCR.

Studying the Expression of Interacting Groups of Genes

A major goal of biologists is to learn how genes act together to produce and maintain a functioning organism. Now that the genomes of a number of species have been sequenced, it is possible to study the expression of large groups of genes—the so-called *systems approach*. Researchers use what is known about the whole genome to investigate which genes are transcribed in different tissues or at different stages of development. One aim is to identify networks of gene expression across an entire genome.

Genome-wide expression studies can be carried out using **DNA microarray assays**. A DNA microarray consists of tiny amounts of a large number of single-stranded DNA fragments representing different genes fixed to a glass slide in a tightly spaced array, or grid, of dots. (The microarray is also called a *DNA chip* by analogy to a computer chip.) Ideally, these fragments represent all the genes of an organism. The mRNAs from cells under study are reverse-transcribed into cDNAs (see Figure 19.10), and a fluorescent label is added so the cDNAs can be used as probes on the microarray. Different fluorescent labels are used for different cell samples so that multiple samples can be tested in the same experiment. The resulting pattern of colored dots, shown in an actual-size microarray in **Figure 19.12**, reveals the dots to which each probe was bound and thus the genes that are expressed in the cell samples being tested. Microarray technology started taking off after several papers about it were published in 1995;

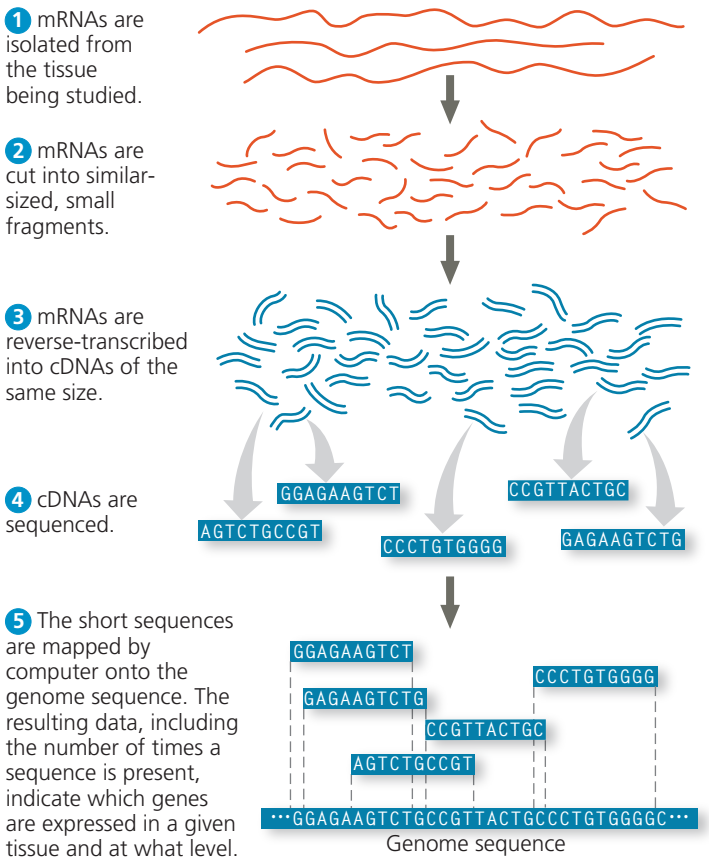
Figure 19.12 Use of microarrays to analyze expression of many genes. In this DNA microarray assay, researchers extracted mRNAs from two different human tissues and synthesized two sets of cDNAs, fluorescently labeled red (tissue #1) or green (tissue #2). Labeled cDNAs were hybridized with a microarray containing 5,760 human genes (about 25% of human genes), part of which is shown in the enlargement. Red indicates that the gene in that well was expressed in tissue #1, green in tissue #2, yellow in both, and black in neither. The fluorescence intensity at each spot indicates the relative expression of the gene.



since then more sophisticated applications have been developed and are in use.

Increasingly, with the advent of rapid, inexpensive DNA sequencing methods, microarray usage is decreasing: Researchers can now afford to simply sequence the cDNA samples from different tissues or different embryonic stages in order to discover which genes are expressed. This straightforward method is called **RNA sequencing**, or **RNA-seq** (pronounced "RNA-seek"), even though it is the cDNA that is actually sequenced. In RNA-seq, the mRNA (or other RNA) samples are isolated, cut into shorter, similar-sized fragments, and converted into cDNAs **(Figure 19.13)**. These short cDNA stretches are sequenced, and a computer program reassembles them, either mapping them onto the genome of the species in question (when available) or simply ordering them from scratch based on overlapping sequences of multiple RNAs. RNA-seq has several advantages over microarrays. First, the procedure is not based on hybridization with a labeled probe, so it doesn't depend on having genomic sequences in hand (although they are usually available). Second, it can measure levels of expression over a very wide range, unlike microarrays, which cannot accurately measure either very low or very high levels. Third, a careful analysis provides a wealth of information about expression of a particular gene, such as relative levels of alternatively spliced mRNAs. As the price of DNA sequencing plummets, RNA-seq is becoming more widely used for many applications. In most cases, however, expression of individual genes still needs to be confirmed by RT-PCR.

Figure 19.13 Use of RNA sequencing (RNA-seq) to analyze expression of many genes. RNA-seq yields a wide range of information about expression of genes, including their level of expression.



Scientists can now measure the expression of thousands of genes at one time. DNA technology makes such studies possible; with automation, they are easily performed on a large scale. By uncovering gene interactions and providing clues to gene function, DNA microarray assays and RNA-seq may contribute to a better understanding of diseases and suggest new diagnostic techniques or therapies. For instance, comparing patterns of gene expression in breast cancer tumors and noncancerous breast tissue has already resulted in more informed and effective treatment protocols (see Figure 18.27). Ultimately, information from these methods should provide a grander view of how ensembles of genes interact to form an organism and maintain its vital systems.

Determining Gene Function

Once they identify a gene of interest, how do scientists determine its function? A gene's sequence can be compared with sequences in other species. If the function of a similar gene in another species is known, one might suspect that the gene product in question performs a comparable task. Data about the location and timing of gene expression may reinforce the suggested function. To obtain stronger evidence, one approach is to disable the gene and then observe the consequences in the cell or organism.

Editing Genes and Genomes

Molecular biologists have long sought techniques for altering, or editing, the genetic material of cells or organisms in a predictable way. In one such technique, called ***in vitro* mutagenesis**, specific mutations are introduced into a cloned gene, and the mutated gene is returned to a cell in such a way that it disables ("knocks out") the normal cellular copies of the same gene. If the introduced mutations alter or destroy the function of the gene product, the phenotype of the mutant cell may help reveal the function of the missing normal protein. Using molecular and genetic techniques worked out in the 1980s, researchers can generate mice with any given gene disabled in order to study the role of that gene in development and in the adult. Mario Capecchi, Martin Evans, and Oliver Smithies received the Nobel Prize in 2007 for developing this technique.

Over the past 10 years, biologists have developed a powerful new technique for gene editing in living cells and organisms, called the **CRISPR-Cas9 system**, that is taking the field of genetic engineering by storm. Cas9 is a bacterial protein that helps defend bacteria against bacteriophage infections in a system worked out by Jennifer Doudna and Emmanuelle Charpentier. In bacterial cells, Cas9 acts

together with a "guide RNA" made from the CRISPR region of the bacterial system (see Figure 26.7).

Similar to the restriction enzymes described earlier, Cas9 is a nuclease that cuts double-stranded DNA molecules. However, while a given restriction enzyme recognizes only one particular DNA sequence, the Cas9 protein will cut any sequence to which it is directed. Cas9 takes its marching orders from a guide RNA molecule that it binds and uses as a homing device, cutting both strands of any DNA sequence that is exactly complementary to the guide RNA. Scientists have been able to exploit the function of Cas9 by introducing a Cas9–guide RNA complex into a cell they wish to alter **(Figure 19.14)**. The guide RNA in the complex is engineered to be complementary to the "target" gene. Cas9 cuts both strands of the target DNA, and the resulting broken ends of DNA trigger a DNA repair system (similar to that shown in Figure 16.19). When there is no undamaged DNA for the enzymes of the repair system to use as a template, as shown at the bottom left of Figure 19.14, the repair enzymes rejoin the ends, sometimes introducing or removing nucleotides. If the cut is directed to a coding portion of the gene, the rejoining process often alters the DNA sequence so that the gene no longer works properly.

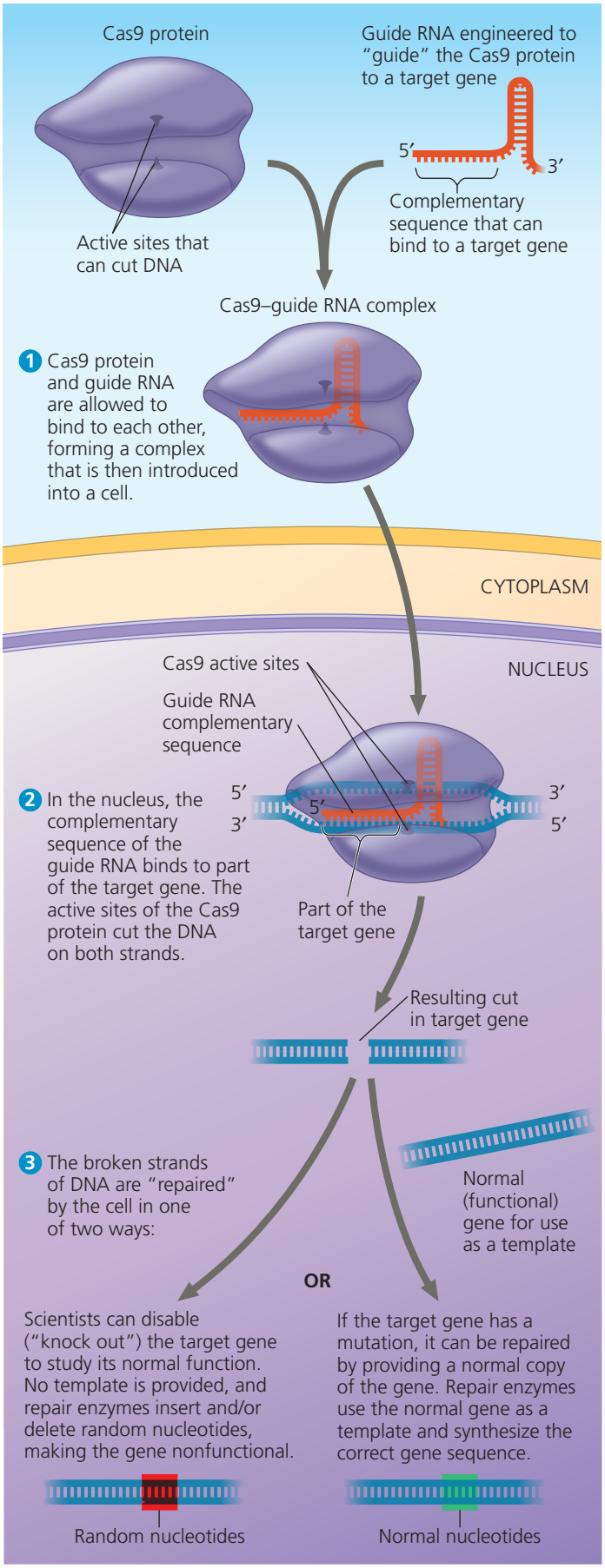
This technique is a highly effective way for researchers to knock out a given gene in order to study what that gene does, and it has already been used in many organisms, including bacteria, fish, mice, insects, human cells, and various crop plants. Researchers have also modified the technique so that CRISPR-Cas9 can be used to repair a gene that has a mutation (see bottom right of Figure 19.14). They introduce a segment from the normal (functional) gene along with the CRISPR-Cas9 system. After Cas9 cuts the target DNA, repair enzymes can use the normal DNA segment as a template to repair the target DNA at the break point. This approach is used for gene therapy, which will be discussed later in the chapter.

In another application of CRISPR-Cas9, scientists are attempting to address the global problem of insect-borne diseases by altering genes in the insect so that, for example, it cannot transmit disease. An extra twist to this approach is engineering the new allele so that it is much more highly favored for inheritance than is the wild-type allele. This is called a **gene drive** because the biased inheritance of the engineered gene during reproduction rapidly "drives" the new allele through the population.

Other Methods for Studying Gene Function

Another method for silencing expression of selected genes doesn't alter the genome; instead, it exploits the phenomenon of **RNA interference (RNAi)**, described in Concept 18.3. This experimental approach uses synthetic double-stranded RNA molecules matching the sequence of a particular gene to trigger breakdown of the gene's messenger RNA or to block its translation. In organisms such as the nematode

Figure 19.14 Gene editing using the CRISPR-Cas9 system.



and the fruit fly, RNAi has already proved valuable for analyzing the functions of genes on a large scale. This method is quicker than using the CRISPR-Cas9 system, but it only leads to a temporary reduction of gene expression rather than a permanent gene knockout or alteration.

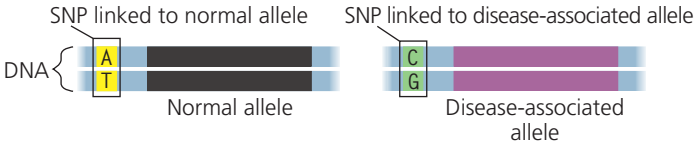
In humans, ethical considerations prohibit knocking out genes to determine their functions. An alternative approach is to analyze the genomes of large numbers of people with a certain phenotypic condition or disease, such as heart disease or diabetes, to try to find differences they all share compared with people without that condition. The assumption is that these differences may be associated with one or more malfunctioning genes, thus in a sense being naturally occurring gene knockouts. In these large-scale analyses, called **genome-wide association studies**, researchers look for *genetic markers*, DNA sequences that vary in the population. In a gene, such sequence variation is the basis of different alleles, as we have seen for sickle-cell disease (see Figure 17.26). And just like the coding sequences of genes, noncoding DNA at a specific locus on a chromosome may exhibit small nucleotide differences among individuals. Variations in coding or noncoding DNA sequences among a population are called polymorphisms (from the Greek for "many forms").

Among the most useful genetic markers in tracking down genes that contribute to diseases and disorders are single base-pair variations in the genomes of the human population. A single base-pair site where variation is found in at least 1% of the population is called a **single nucleotide polymorphism** (**SNP**, pronounced "snip"). A few million SNPs occur in the human genome, about once in 100–300 base pairs of both coding and noncoding DNA sequences. To find SNPs in large numbers of people, it isn't necessary to sequence their DNA; SNPs can be detected by very sensitive microarray assays, RNA-seq, or PCR.

Once a SNP is identified that is found in all affected people, researchers focus on that region and sequence it. In nearly all cases, the SNP itself does not contribute directly to the disease in question by altering the encoded protein; in fact, most SNPs are in noncoding regions. Instead, if the SNP and a disease-associated allele are close enough to be genetically linked, scientists can take advantage of the fact that crossing over between the marker and the gene is very unlikely during gamete formation. Therefore, the marker and gene will almost always be inherited together, even though the marker is not part of the gene **(Figure 19.15)**. SNPs have been found that correlate with diabetes, heart disease, and several types of cancer, and the search is on for genes that might be involved.

The experimental approaches you have learned about thus far focused on working with molecules, mainly DNA and proteins. In a parallel line of inquiry, biologists have been developing powerful techniques for cloning whole multicellular organisms. One aim of this work is to obtain special types of cells, called stem cells, that can give rise to all types of tissues. Being able to manipulate stem cells would allow scientists to use the DNA-based methods previously discussed to alter stem cells for the treatment of diseases. Methods involving the cloning of organisms and production of stem cells are the subject of the next section.

Figure 19.15 Single nucleotide polymorphisms (SNPs) as genetic markers for disease-associated alleles. This diagram depicts the same region of the genome from two groups of individuals, one group having a particular disease or condition with a genetic basis. Unaffected people have an A/T pair at a given SNP locus, while affected people have a C/G pair there. Once the allele is confirmed as being associated with the disease in question, the SNP that varies in this way can be used as a marker for the disease-associated allele.



CONCEPT CHECK 19.2

1. Describe the role of complementary base pairing during RT-PCR, DNA microarray analysis, RNA sequencing, and CRISPR-Cas9 editing.
2. **VISUAL SKILLS ➤** Consider the microarray in Figure 19.12. If a sample from normal tissue is labeled with a green fluorescent dye and a sample from cancerous tissue is labeled red, what color spots would represent genes you would be interested in if you were studying cancer? Explain.

For suggested answers, see Appendix A.

CONCEPT 19.3

Cloned organisms and stem cells are useful for basic research and other applications

Along with advances in DNA technology, scientists have been developing and refining methods for cloning whole multicellular organisms from single cells. In this context, cloning produces one or more organisms that are genetically identical to the "parent" that donated the single cell. This is often called *organismal cloning* to differentiate it from gene cloning and, more significantly, from cell cloning—the division of an asexually reproducing cell such as a bacterium into a group of genetically identical cells. (The common theme is that the product is genetically identical to the parent. In fact, the word *clone* comes from the Greek *klon*, meaning "twig.") The current interest in organismal cloning arises primarily from its ability to generate stem cells. A **stem cell** is a relatively

unspecialized cell that can both reproduce itself indefinitely and, under appropriate conditions, differentiate into specialized cells of one or more types. Stem cells have great potential for regenerating damaged tissues.

The cloning of plants and animals was first attempted over 50 years ago in experiments designed to answer basic biological questions. For example, researchers wondered if all the cells of an organism have the same genes or whether cells lose genes during the process of differentiation (see Concept 18.4). One way to answer this question is to see whether a differentiated cell can generate a whole organism—in other words, whether cloning an organism is possible. Let's discuss these early experiments before we consider more recent progress in organismal cloning and procedures for producing stem cells.

Cloning Plants: Single-Cell Cultures

The successful cloning of whole plants from single differentiated cells was accomplished during the 1950s by F. C. Steward and his students at Cornell University, who worked with carrot plants. They found that differentiated cells taken from the root (the carrot) and incubated in culture medium could grow into normal adult plants, each genetically identical to the parent plant. These results showed that differentiation does not necessarily involve irreversible changes in the DNA. In plants, mature cells can "dedifferentiate" and then give rise to all the specialized cell types of the organism. Any cell with this potential is said to be **totipotent**.

Plant cloning is used extensively in agriculture. For plants such as orchids, cloning is the only commercially practical means of reproducing plants. In other cases, cloning has been used to reproduce a plant with valuable characteristics, such as resistance to plant pathogens. In fact, you yourself may be a plant cloner: If you have ever grown a new plant from a cutting, you have practiced cloning!

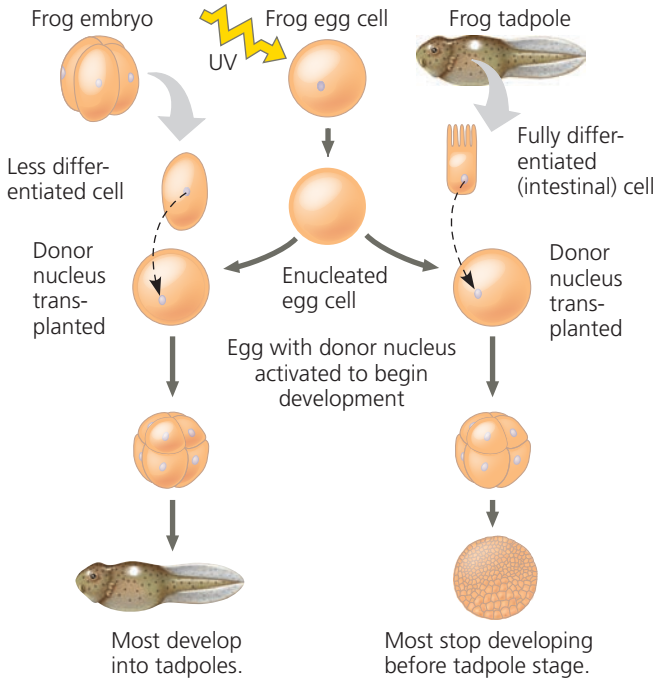
Cloning Animals: Nuclear Transplantation

Differentiated cells from animals generally do not divide in culture, much less develop into the multiple cell types of a new organism. Therefore, early researchers had to use a different approach to answer the question of whether differentiated animal cells are totipotent. Their approach was to remove the nucleus of an egg (creating an *enucleated* egg) and replace it with the nucleus of a differentiated cell, a procedure called *nuclear transplantation*, now more commonly called *somatic cell nuclear transfer*. If the nucleus from the differentiated donor cell retains its full genetic capability, then it should be able to direct development of the recipient cell into all the tissues and organs of an organism. Such experiments were conducted on one species of frog (*Rana pipiens*) by Robert Briggs and Thomas King in the 1950s and on another frog species (*Xenopus laevis*) by John Gurdon in the 1970s **(Figure 19.16)**. These researchers transplanted a nucleus from an embryonic or tadpole cell into an enucleated egg of the same species. In

Figure 19.16

Inquiry Can the nucleus from a differentiated animal cell direct development of an organism?

Experiment John Gurdon and colleagues at Oxford University, in England, destroyed the nuclei of frog (*Xenopus laevis*) eggs by exposing the eggs to ultraviolet light. They then transplanted nuclei from cells of frog embryos and tadpoles into the enucleated eggs.



Results When the transplanted nuclei came from an early embryo, the cells of which are relatively undifferentiated, most of the recipient eggs developed into tadpoles. But when the nuclei came from the fully differentiated intestinal cells of a tadpole, fewer than 2% of the eggs developed into normal tadpoles, and most of the embryos stopped developing at a much earlier stage.

Conclusion The nucleus from a differentiated frog cell can direct development of a tadpole. However, its ability to do so decreases as the donor cell becomes more differentiated, presumably because of changes in the nucleus.

Data from J. B. Gurdon et al., The developmental capacity of nuclei transplanted from keratinized cells of adult frogs, *Journal of Embryology and Experimental Morphology* 34:93–112 (1975).

WHAT IF? ➤ *If each cell in a four-cell embryo were already so specialized that it was not totipotent, what results would you predict for the experiment on the left side of the figure?*

HHMI Video: Somatic Cell Nuclear Transfer

Gurdon's experiments, the transplanted nucleus was often able to support normal development of the egg into a tadpole. However, he found that the potential of a transplanted nucleus to direct normal development was inversely related to the age of the donor: The older the donor nucleus, the lower the percentage of normal tadpoles (see Figure 19.16).

From these results, Gurdon concluded that something in the nucleus *does* change as animal cells differentiate. In

➤ Figure 19.17 Reproductive cloning of a mammal by nuclear transfer. Dolly, shown here as a lamb, has a very different appearance from her surrogate mother, standing beside her.

frogs and most other animals, nuclear potential tends to be restricted more and more as embryonic development and cell differentiation progress. These were foundational experiments that ultimately led to stem cell technology, and Gurdon received the 2012 Nobel Prize in Medicine for this work.

Reproductive Cloning of Mammals

In addition to cloning frogs, researchers were able to clone mammals using early embryonic cells as a source of donor nuclei. Until about 20 years ago, though, it was not known whether a nucleus from a fully differentiated cell could be reprogrammed successfully to act as a donor nucleus. In 1997, researchers in Scotland announced the birth of Dolly, a lamb cloned from an adult sheep by nuclear transfer from a differentiated mammary gland cell **(Figure 19.17)**. Using a technique related to that in 19.16, the researchers implanted early embryos into surrogate mothers. Out of several hundred embryos, one successfully completed normal development, and Dolly was born, a genetic clone of the nucleus donor. At the age of 6, Dolly suffered complications from a lung infection often seen in sheep kept indoors and was euthanized. Another cloned sheep from the same experiment developed an unusual lung disease. This led to speculation that this sheep's cells were in some way not quite as healthy as those of a normal sheep, possibly reflecting incomplete reprogramming of the original transplanted nucleus. Reprogramming involves epigenetic changes that lead to changes in chromatin structure (see Concept 18.2), to be discussed shortly.

Since that time, researchers have cloned numerous other mammals, including mice, cats, cows, horses, pigs, dogs, and monkeys. In most cases, their goal has been the production of new individuals; this is known as *reproductive cloning*. We have already learned a lot from such experiments. For example, cloned animals of the same species do *not* always look or behave identically. In a herd of cows cloned from the same line of cultured cells, certain cows are dominant in behavior and others are more submissive. Another example of nonidentity in clones is the first cloned cat, named CC for Carbon Copy **(Figure 19.18)**. She has a calico coat, like her single

➤ Figure 19.18 CC ("Carbon Copy"), the first cloned cat (right), and her single parent. Rainbow (left) donated the nucleus in a cloning procedure that resulted in CC. However, the two cats are not identical: Rainbow is a classic calico cat with orange patches on her fur and has a "reserved personality," while CC has a gray and white coat and is more playful.

female parent, but the color and pattern are different because of random X chromosome inactivation, which is a normal occurrence during embryonic development (see Figure 15.8). And identical human twins, which are naturally occurring "clones," are always slightly different. Clearly, environmental influences and random phenomena play a significant role during development.

Faulty Gene Regulation in Cloned Animals Due to Epigenetic Differences

In most nuclear transplantation studies thus far, only a small percentage of cloned embryos develop normally to birth. And like Dolly, many cloned animals exhibit defects. Cloned mice, for instance, are prone to obesity, pneumonia, liver failure, and premature death. Scientists assert that even cloned animals that appear normal are likely to have subtle defects.

Researchers have uncovered some reasons for the low efficiency of cloning and the high incidence of abnormalities. In the nuclei of fully differentiated cells, a small subset of genes is turned on and expression of the rest of the genes is repressed. This regulation often is the result of epigenetic changes in chromatin, such as acetylation of histones or methylation of DNA (see Figure 18.7). During the nuclear transfer procedure, many of these changes must be reversed in the later-stage nucleus from a donor animal for genes to be expressed or repressed appropriately in early stages of development. Researchers have found that the DNA in cells from cloned embryos, like that of differentiated cells, often has more methyl groups than does the DNA in equivalent cells from normal embryos of the same species. This finding suggests that the reprogramming of donor nuclei requires more accurate

and complete chromatin restructuring than occurs during cloning procedures. Because DNA methylation helps regulate gene expression, misplaced or extra methyl groups in the DNA of donor nuclei may interfere with the pattern of gene expression necessary for normal embryonic development. In fact, the success of a cloning attempt may depend in large part on whether or not the chromatin in the donor nucleus can be artificially modified to resemble that of a newly fertilized egg.

Stem Cells of Animals

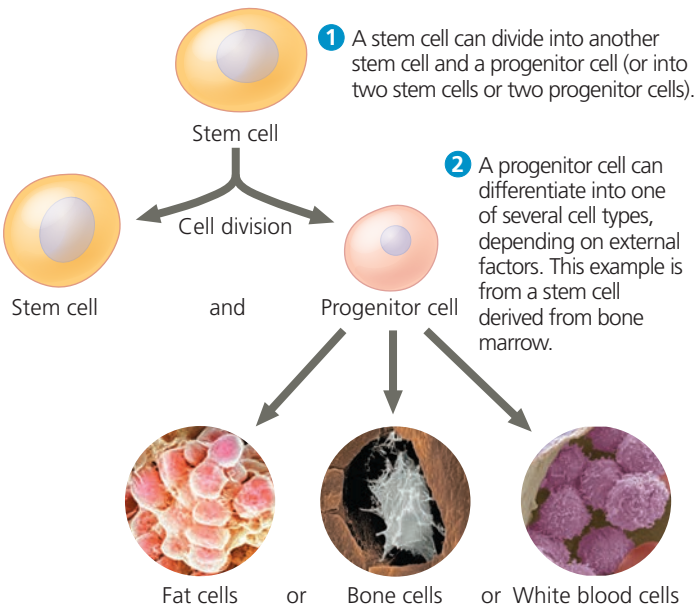
Progress in cloning mammalian embryos, including primates, has heightened speculation about the cloning of humans, which has not yet been achieved past very early embryonic stages. The main reason researchers have been trying to clone human embryos is not for reproduction, but for the production of stem cells to treat human diseases. Recall that a stem cell is a relatively unspecialized cell that can both reproduce itself indefinitely and, under appropriate conditions, differentiate into specialized cells of one or more types **(Figure 19.19)**. Thus, stem cells are able to both replenish their own population and generate cells that travel down specific differentiation pathways.

HHMI Animation: Somatic Cell Nuclear Transfer

Embryonic and Adult Stem Cells

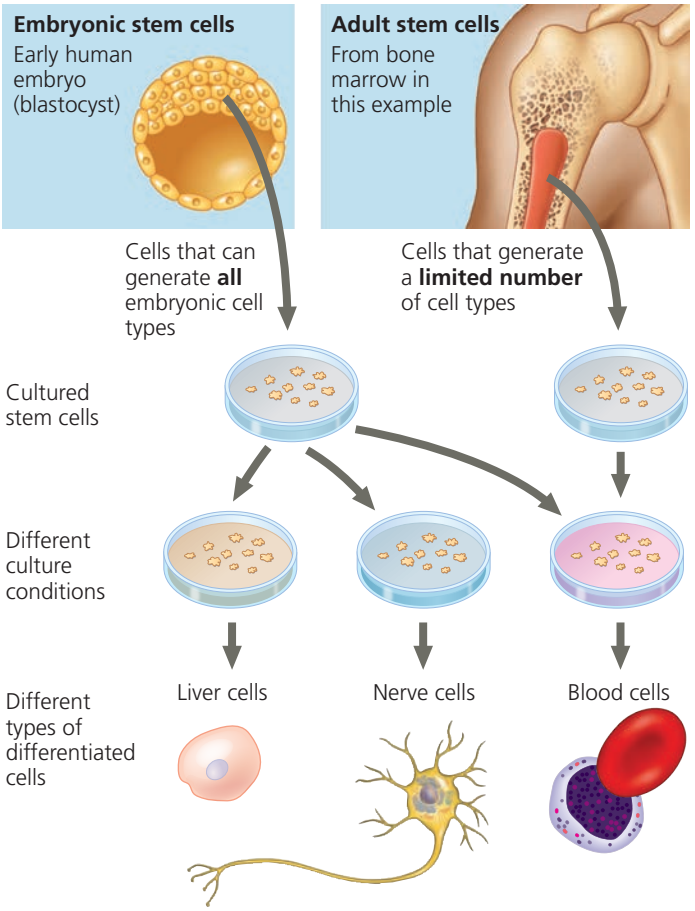
Many early animal embryos contain stem cells capable of giving rise to differentiated cells of any type. Stem cells can be isolated from early embryos at a stage called the blastula stage or its human equivalent, the blastocyst stage. In culture, these *embryonic stem (ES) cells* reproduce indefinitely, and depending on culture conditions, they can be made to differentiate

Figure 19.19 How stem cells maintain their own population and generate differentiated cells.



HHMI Video: Cultured Human Embryonic Stem Cells

Figure 19.20 Working with stem cells. Animal stem cells, which can be isolated from early embryos or adult tissues and grown in culture, are self-perpetuating, relatively undifferentiated cells. Embryonic stem cells are easier to grow than adult stem cells and can theoretically give rise to *all* types of cells in an organism. The range of cell types that can arise from adult stem cells is not yet fully understood.



HHMI Animation: Creating Embryonic Stem Cell Lines

into a wide variety of specialized cells **(Figure 19.20)**, including even eggs and sperm.

The adult body also has stem cells, which serve to replace nonreproducing specialized cells as needed. In contrast to ES cells, *adult stem cells* are not able to give rise to all cell types in the organism, though they can generate several defined types. For example, one of the several types of stem cells in bone marrow can generate all the different kinds of blood cells (see Figure 19.20), and another type of bone marrow stem cell can differentiate into bone, cartilage, fat, muscle, and the linings of blood vessels. To the surprise of many, the adult brain has been found to contain stem cells that continue to produce certain kinds of nerve cells there. Researchers have also reported finding stem cells in skin, hair, eyes, and dental pulp. Although adult animals have only tiny numbers of stem cells, scientists are learning to identify and isolate these cells from various tissues and, in some cases, to grow them in culture. With the right culture conditions (for instance, the addition of specific growth factors), cultured stem cells from adult animals have been made to differentiate

into various defined types of specialized cells, although none are as versatile as ES cells.

Research with embryonic or adult stem cells is a source of valuable data about differentiation and has enormous potential for medical applications. The ultimate aim is to supply cells for the repair of damaged or diseased organs: for example, insulin-producing pancreatic cells for people with type 1 diabetes or certain kinds of brain cells for people with Parkinson's disease or Huntington's disease. Adult stem cells from bone marrow have long been used in bone marrow transplants as a source of immune system cells in patients whose own immune systems are nonfunctional because of genetic disorders or radiation treatments for cancer.

The developmental potential of adult stem cells is limited to certain tissues. ES cells hold more promise than adult stem cells for most medical applications because ES cells are **pluripotent**, capable of differentiating into many different cell types. In 2013, a research group reported that they had established ES cell lines from human blastocysts produced by transferring a nucleus from a differentiated cell into an enucleated egg. Prior to that report, cells were obtained only from embryos donated by patients undergoing infertility treatments or from long-term cell cultures originally established with cells isolated from donated embryos, an issue that prompts ethical and political discussions. Although the techniques for cloning early human embryos are still being optimized, they represent a potential new source for ES cells that may be less controversial. Furthermore, with a donor nucleus from a person with a particular disease, researchers should be able to produce ES cells that match the patient and are thus not rejected by his or her immune system when used for treatment. When the main aim of cloning is to produce ES cells to treat disease, the process is called *therapeutic cloning*. Although most people believe that reproductive cloning of humans is unethical, opinions vary about the morality of therapeutic cloning.

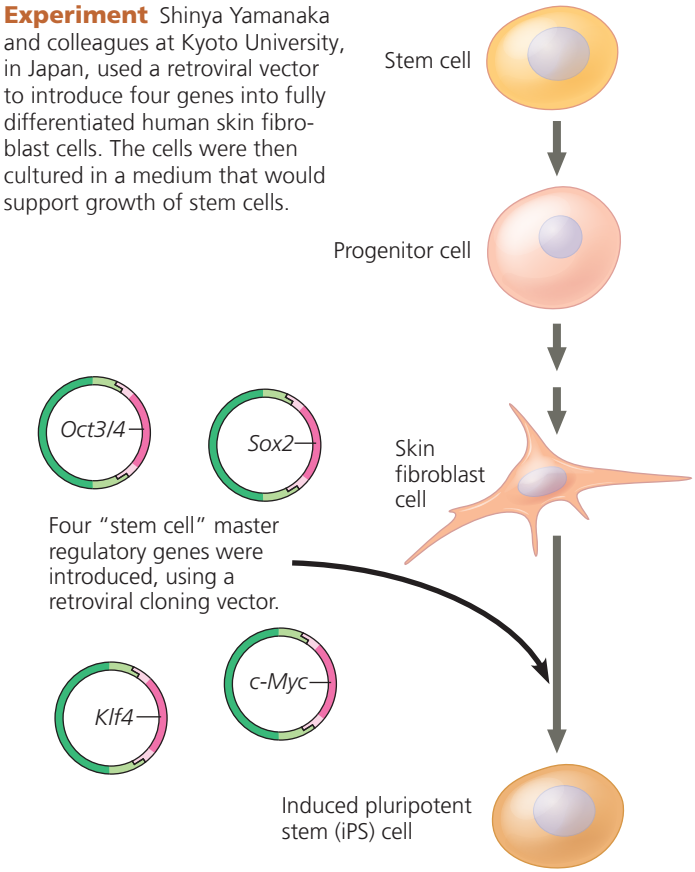
Induced Pluripotent Stem (iPS) Cells

Resolving the debate now seems less urgent because researchers have learned to turn back the clock in fully differentiated cells, reprogramming them to act like ES cells. The accomplishment of this feat, which posed formidable obstacles, was announced in 2007, first by labs using mouse skin cells and then by additional groups using cells from human skin and other organs or tissues. In all these cases, researchers transformed the differentiated cells into a type of ES cell by using a retrovirus (a specific class of virus; see Table 26.1) to introduce extra, cloned copies of four "stem cell" master regulatory genes. The "deprogrammed" cells are known as *induced pluripotent stem* (*iPS*) cells because, in using this fairly simple laboratory technique to return them to their undifferentiated state, pluripotency has been restored. The experiments that first transformed human differentiated cells into iPS cells are described in **Figure 19.21**. Shinya Yamanaka received the 2012 Nobel Prize in Medicine for this work, shared with John Gurdon, whose work you read about in Figure 19.16.

By many criteria, iPS cells can perform most of the functions of ES cells, but there are some differences in gene expression and other cellular functions, such as cell division. At least until these differences are fully understood, the study

Figure 19.21

Inquiry Can a fully differentiated human cell be "deprogrammed" to become a stem cell?



Experiment Shinya Yamanaka and colleagues at Kyoto University, in Japan, used a retroviral vector to introduce four genes into fully differentiated human skin fibroblast cells. The cells were then cultured in a medium that would support growth of stem cells.

Results Two weeks later, the cells resembled embryonic stem cells in appearance and were actively dividing. Their gene expression patterns, gene methylation patterns, and other characteristics were also consistent with those of embryonic stem cells. The iPS cells were able to differentiate into heart muscle cells, as well as other cell types.

Conclusion The four genes induced differentiated skin cells to become pluripotent stem cells, with characteristics of embryonic stem cells.

Data from K. Takahashi et al., Induction of pluripotent stem cells from adult human fibroblasts by defined factors, *Cell* 131:861–872 (2007).

WHAT IF? ➤ *Patients with diseases such as heart disease or Alzheimer's could have their own skin cells reprogrammed to become iPS cells. Once procedures have been developed for converting iPS cells into heart or nervous system cells, the patients' own iPS cells might be used to treat their disease. When organs are transplanted from a donor to a diseased recipient, the recipient's immune system may reject the transplant, a dangerous condition. Would using iPS cells be expected to carry the same risk? Why or why not? Given that these cells are actively dividing, undifferentiated cells, what risks might this procedure carry?*

of ES cells will continue to make important contributions to the development of stem cell therapies. (In fact, it is likely that ES cells will always be a focus of basic research as well.) In the meantime, work is proceeding using the iPS cells that have been experimentally produced.

There are two major potential uses for human iPS cells. First, cells from patients suffering from diseases have been reprogrammed to become iPS cells, which act as model cells for studying the disease and potential treatments. Human iPS cell lines have already been developed from individuals with type 1 diabetes, Parkinson's disease, Huntington's disease, Down syndrome, and many other diseases. Second, in the field of regenerative medicine, a patient's own cells could be reprogrammed into iPS cells and then used to replace nonfunctional tissues, such as insulin-producing cells of the pancreas. In fact, in 2014 two research groups described successful methods for growing insulin-producing cells from both iPS cells and ES cells. Before these cells can be used in patients, however, researchers will have to develop a way to ensure the cells are not destroyed by the patients' immune system (the original cause of type 1 diabetes, in which the immune system malfunctions).

In another surprising development, researchers have been able to identify genes that can directly reprogram a differentiated cell into another type of differentiated cell without it passing through a pluripotent state. In the first reported example, one type of cell in the pancreas was transformed into another type. However, the two types of cells do not need to be very closely related: Another research group has been able to directly reprogram a skin fibroblast into a nerve cell. Development techniques that direct iPS cells or even fully differentiated cells to become specific cell types for regenerative medicine is an area of intense research, one that has already seen some success. The iPS cells created in this way could eventually provide tailor-made "replacement" cells for patients without using any human eggs or embryos, thus circumventing most ethical objections.

CONCEPT CHECK 19.3

1. Based on current knowledge, how would you explain the difference in the percentage of tadpoles that developed from the two kinds of donor nuclei in Figure 19.16?
2. A few companies in China and South Korea provide the service of cloning dogs, using cells from their clients' pets to provide nuclei in procedures like that in Figure 19.17. Should their clients expect the clone to look identical to their original pet? Why or why not? What ethical questions does this bring up?
3. **MAKE CONNECTIONS ➤** Based on what you know about muscle differentiation (see Figure 18.18) and genetic engineering, propose the first experiment you might try if you wanted to direct an embryonic stem cell or iPS cell to develop into a muscle cell.

For suggested answers, see Appendix A.

CONCEPT 19.4

The practical applications of DNA-based biotechnology affect our lives in many ways

DNA technology is in the news almost every day. Most often, the topic is a new and promising application in medicine, but this is just one of numerous fields benefiting from DNA technology and genetic engineering.

Medical Applications

One important use of DNA technology is the identification of human genes whose mutation plays a role in genetic diseases. These discoveries may lead to ways of diagnosing, treating, and even preventing such conditions. DNA technology is also contributing to our understanding of "nongenetic" diseases, from arthritis to AIDS, since a person's genes influence susceptibility to these diseases. Furthermore, diseases of all sorts involve changes in gene expression within the affected cells and often within the patient's immune system. By using RNA-seq and DNA microarray assays (see Figures 19.12 and 19.13) or other techniques to compare gene expression in healthy and diseased tissues, researchers are finding genes that are turned on or off in particular diseases. These genes and their products are potential targets for prevention or therapy.

Diagnosis and Treatment of Diseases

A new chapter in the diagnosis of infectious diseases has been opened by DNA technology, in particular the use of PCR and labeled nucleic acid probes to track down pathogens. For example, because the sequence of the RNA genome of HIV is known, RT-PCR can be used to amplify, and thus detect, HIV RNA in blood or tissue samples (see Figure 19.11). RT-PCR is often the best way to detect an otherwise elusive infective agent.

Medical scientists can now diagnose hundreds of human genetic disorders by using PCR with primers that target the genes associated with these disorders. The amplified DNA product is then sequenced to reveal the presence or absence of the disease-causing mutation. Among the genes for human diseases that have been identified are those for sickle-cell disease, hemophilia, cystic fibrosis, Huntington's disease, and Duchenne muscular dystrophy. Individuals with such diseases can often be identified before the onset of symptoms, even before birth (see Figure 14.19). PCR can also be used to identify symptomless carriers of potentially harmful recessive alleles.

As you learned earlier, genome-wide association studies have pinpointed SNPs (single nucleotide polymorphisms) that are linked to disease-associated alleles (see Figure 19.15).

Individuals can be tested by PCR and sequencing for a SNP that is correlated with the abnormal allele. The presence of particular SNPs is correlated with increased risk for conditions such as heart disease, Alzheimer's disease, and some types of cancer. Companies that offer individual genetic testing for risk factors like these are looking for previously identified, linked SNPs. It may be helpful for individuals to learn about their health risks, with the understanding that such genetic tests merely reflect correlations and do not make predictions.

The techniques described in this chapter have also prompted improvements in disease treatments. By analyzing the expression of many genes in breast cancer patients, researchers have been able to refine their understanding of the different subtypes of breast cancer (see Figure 18.27). Knowing the expression levels of particular genes can help physicians determine the likelihood that the cancer will recur, thus helping them design an appropriate treatment. Given that some low-risk patients have a 96% survival rate over a ten-year period with no treatment, gene expression analysis allows doctors and patients access to valuable information when they are considering treatment options.

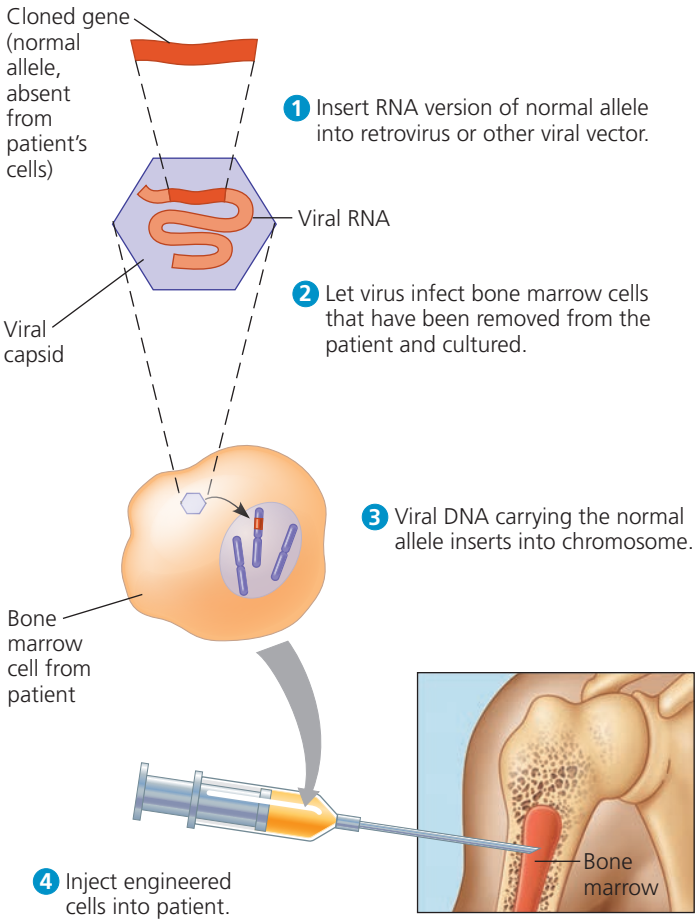
Many envision a future of "personalized medicine" where each person's genetic profile can inform them about diseases or conditions for which they are especially at risk and help them make treatment choices. As we will discuss later in the chapter, a *genetic profile* is currently taken to mean a set of genetic markers such as SNPs. Ultimately, however, it will likely mean the complete DNA sequence of an individual—once sequencing becomes inexpensive enough. Our ability to sequence a person's genome rapidly and inexpensively is advancing faster than our development of appropriate treatments for the conditions we are characterizing. Still, the identification of genes involved in these conditions provides us with good targets for therapeutic interventions.

Human Gene Therapy and Gene Editing

Gene therapy—the introduction of genes into an afflicted individual for therapeutic purposes—holds great potential for treating the relatively small number of disorders traceable to a single defective gene. The aim of this approach is to insert a normal allele of the defective gene into the somatic cells of the tissue affected by the disorder.

For gene therapy of somatic cells to be permanent, the cells that have the normal allele must be cells that multiply throughout the patient's life. Bone marrow cells, which include the stem cells that give rise to all the cells of the blood and immune system, are prime candidates. **Figure 19.22** outlines one procedure for gene therapy in an individual whose bone marrow cells do not produce a vital

Figure 19.22 Gene therapy using a retroviral vector. A retrovirus that has been rendered harmless is used as a vector in this procedure, which exploits the ability of a retrovirus to insert a DNA transcript of its RNA genome into the chromosomal DNA of its host cell (see Figure 26.9). If the foreign gene carried by the retroviral vector is expressed, the cell and its descendants will possess the gene product. Cells that reproduce throughout life, such as bone marrow cells, are ideal candidates for gene therapy.



enzyme because of a single defective gene. One type of severe combined immunodeficiency (SCID) is caused by this kind of defect. If the treatment is successful, the patient's bone marrow cells will begin producing the missing protein, and the patient may be cured.

The procedure shown in Figure 19.22 was used in gene therapy trials for SCID in France in 2000. In that trial, ten young children with SCID were treated by the same procedure. Nine of these patients showed significant, definitive improvement after two years, the first indisputable success of gene therapy. However, three of the patients subsequently developed leukemia, a type of blood cell cancer, and one of them died. Researchers have concluded it is likely that the insertion of the retroviral vector occurred near a gene that triggers the proliferation of blood cells. Using a viral vector that does not come from a retrovirus, clinical researchers have treated at least three other genetic diseases somewhat

successfully with gene therapy: a type of progressive blindness (see Concept 50.3), a degenerative disease of the nervous system, and a blood disorder involving the β-globin gene.

Gene therapy raises many technical issues. For example, how can the activity of the transferred gene be controlled so that cells make appropriate amounts of the gene product at the right time and in the right place? How can we be sure that the insertion of the therapeutic gene does not harm some other necessary cell function? As more is learned about DNA control elements and gene interactions, researchers may be able to answer such questions.

A more direct approach that avoids the complications of using a viral vector in gene therapy is made possible by gene editing, especially given the development of the CRISPR-Cas9 system described earlier. In this approach the existing defective gene is edited to correct the mutation. As shown in Figure 19.14, the CRISPR-Cas9 system is capable of doing this.

In 2014, a group of researchers reported correcting a genetic defect in mice using CRISPR-Cas9 technology. The lab mice had been genetically engineered to have a mutation in a gene encoding a liver enzyme that metabolizes the amino acid tyrosine, mimicking a fatal genetic disorder in humans called tyrosinemia. A guide RNA molecule complementary to the mutated region of the gene was introduced into the mouse along with the Cas9 protein and a segment of DNA from the same region of the normal gene for use as a template. Subsequent analysis indicated that the faulty gene had been corrected in enough of the liver cells that the amount of functional enzyme made was sufficient to alleviate the disease symptoms. There are still hurdles to overcome before this approach can be used in clinical trials in humans, but the CRISPR technology is sparking widespread excitement among researchers and physicians alike.

In addition to technical challenges, gene therapy and gene editing provoke ethical questions. Some critics believe that tampering with human genes in any way is immoral or unethical. Other observers see no fundamental difference between the transplantation of genes into somatic cells and the transplantation of organs. You might wonder whether scientists are considering engineering human germ-line cells in the hope of correcting a defect in future generations. Such genetic engineering is now routinely done in laboratory mice, and, in fact, conditions that would allow genetic engineering of human embryos have been worked out.

The development of the CRISPR-Cas9 system has engendered much debate about the ethics of gene editing, related to applications both potential and real. In March of 2015, an editorial was published by leading scientists working with CRISPR-Cas9 calling for the research community to "strongly discourage" any experimental work on human eggs or embryos. A month later, however, scientists in China reported using CRISPR-Cas9 technology to edit a gene in human embryos. (They used "nonviable" fertilized eggs—zygotes—that would not develop all the way but could form blastocysts.) The researchers were attempting to edit the β-thalassemia gene, mutations in which cause a blood disease of the same name. They injected 86 zygotes, only four of which showed the gene to be edited properly. In many of the other embryos, there were effects on genes other than the β-thalassemia gene—at a much higher level than had been seen in mouse embryos or human cell lines. This study highlighted problems with the technique, at least in human embryos, and at the same time accelerated concern about ethical considerations. Under what circumstances, if any, should we alter the genomes of human germ lines? Would this inevitably lead to the practice of eugenics, a deliberate effort to control the genetic makeup of human populations? While we may not have to resolve these questions immediately, considering them is imperative because they will likely come to the fore at some point in the near future.

Pharmaceutical Products

The pharmaceutical industry derives significant benefit from advances in DNA technology and genetic research, applying them to the development of useful drugs to treat diseases. Pharmaceutical products are synthesized using methods of either organic chemistry or biotechnology, depending on the nature of the product.

Synthesis of Small Molecules for Use as Drugs

Determining the sequence and structure of proteins crucial for tumor cell survival has led to the identification of small molecules that combat certain cancers by blocking the function of these proteins. One drug, imatinib (trade name Gleevec), is a small molecule that inhibits one tyrosine kinase (see Figure 9.8). The overexpression of this kinase, resulting from a chromosomal translocation, is instrumental in causing chronic myelogenous leukemia (CML; see Figure 15.16). Patients in the early stages of CML who are treated with imatinib have exhibited nearly complete, sustained remission from the cancer. Drugs that work in a similar way have also been used with success to treat a few types of lung and breast cancers. This approach is feasible only for cancers for which the molecular basis is fairly well understood.

HHMI Video: Gleevec

In many cases of such drug-treated tumors, though, cells later arise that are resistant to the new drug. In one study, the whole genome of the tumor cells was sequenced both before and after the appearance of drug resistance. Comparison of the sequences showed genetic changes that allowed the tumor cells to "get around" the drug-inhibited protein. Here, we can see that cancer cells demonstrate the principles of evolution: Certain tumor cells have a random mutation that allows them to survive in the presence of a particular drug, and as a consequence of natural selection in the presence of the drug, these are the cells that survive and reproduce.

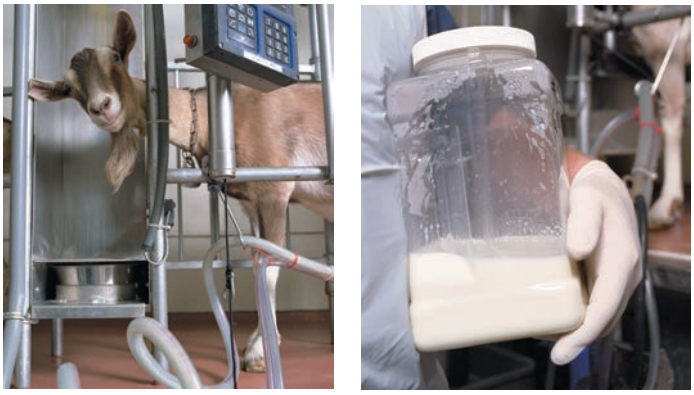
Protein Production in Cell Cultures Pharmaceutical products that are proteins are commonly synthesized on a large scale using cell cultures. You learned earlier in the chapter about DNA cloning and gene expression systems for producing large quantities of a chosen protein that is present naturally in only minute amounts. The host cells used in such expression systems can even be engineered to secrete a protein as it is made, thereby simplifying the task of purifying it by traditional biochemical methods.

Among the first pharmaceutical products manufactured in this way were human insulin and human growth hormone (HGH). Some 2 million people with diabetes in the United States depend on insulin treatment to control their disease. Human growth hormone has been a boon to children born with a form of dwarfism caused by inadequate amounts of HGH, as well as helping AIDS patients gain weight. Another important pharmaceutical product produced by genetic engineering is tissue plasminogen activator (TPA). If administered shortly after a heart attack, TPA helps dissolve blood clots and reduces the risk of subsequent heart attacks.

Protein Production by "Pharm" Animals In some cases, instead of using cell systems to produce large quantities of protein products, pharmaceutical scientists can use whole animals. They can introduce a gene (or other DNA) from an animal into the genome of another individual, often of a different species. This individual is then called a **transgenic** animal. To do this, they first remove eggs from a female of the recipient species and fertilize them *in vitro*. Meanwhile, they have cloned the desired gene from the donor organism. They then inject the cloned DNA directly into the nuclei of the fertilized eggs. Some of the cells integrate the foreign DNA, the *transgene*, into their genome and are able to express the foreign gene. The engineered embryos that arise from these zygotes are then surgically implanted in a surrogate mother. If an embryo develops successfully, the result is a transgenic animal that expresses its new, "foreign" gene.

Assuming that the introduced gene encodes a protein desired in large quantities, these transgenic animals can act as pharmaceutical "factories." For example, a transgene for a human blood protein such as antithrombin, which prevents blood clots, can be inserted into the genome of a goat in such a way that the transgene's product is secreted in the animal's milk **(Figure 19.23)**. The protein is then purified from the milk (which is easier than purification from a cell culture). Such proteins must be tested to ensure that they (or contaminants from the farm animals) will not cause allergic reactions or other adverse effects in patients who receive them.

Figure 19.23 Goats as "pharm" animals. This transgenic goat carries a gene for a human blood protein, antithrombin, which she secretes in her milk. Patients with a rare hereditary disorder in which this protein is lacking suffer from formation of blood clots in their blood vessels. Easily purified from the goat's milk, the protein is used to prevent blood clots in these patients during surgery or childbirth.



Forensic Evidence and Genetic Profiles

In violent crimes, body fluids or small pieces of tissue may be left at the scene or on the clothes or other possessions of the victim or assailant. If enough blood, semen, or tissue is available, forensic laboratories can determine the blood type or tissue type by using antibodies to detect specific cell-surface proteins. However, such tests require fairly fresh samples in relatively large amounts. Also, because many people have the same blood or tissue type, this approach can only exclude a suspect; it cannot provide strong evidence of guilt.

DNA testing, on the other hand, can identify the guilty individual with a high degree of certainty because the DNA sequence of every person is unique (except for identical twins). Genetic markers that vary in the population can be analyzed for a given person to determine that individual's unique set of genetic markers, or **genetic profile**. (This term is preferred over "DNA fingerprint" by forensic scientists, who want to emphasize the heritable aspect of these markers rather than the fact that they produce a pattern on a gel that, like a fingerprint, is visually recognizable.) The FBI started applying DNA technology to forensics in 1988, using a method involving gel electrophoresis and nucleic acid hybridization to detect similarities and differences in DNA samples. This method required much smaller samples of blood or tissue than earlier methods—only about 1,000 cells.

Animation: Genetic Profiles

Today, forensic scientists use an even more sensitive method that takes advantage of variations in length of genetic markers called **short tandem repeats (STRs)**. These are tandemly repeated units of two- to five-nucleotide sequences in specific regions of the genome. The number of repeats present in these regions is highly variable from person to person (polymorphic), and even for a single individual, the two alleles of an STR may differ from each other. For example, one individual may have the sequence ACAT repeated

30 times at one genome locus and 15 times at the same locus on the other homolog, whereas another individual may have 18 repeats at this locus on each homolog. (These two genotypes can be expressed by the two repeat numbers: 30,15 and 18,18.) PCR is used to amplify particular STRs, using sets of primers that are labeled with different-colored fluorescent tags; the length of the region, and thus the number of repeats, can then be determined by electrophoresis. The PCR step allows use of this method even when the DNA is in poor condition or available only in minute quantities: A tissue sample containing as few as 20 cells can be sufficient.

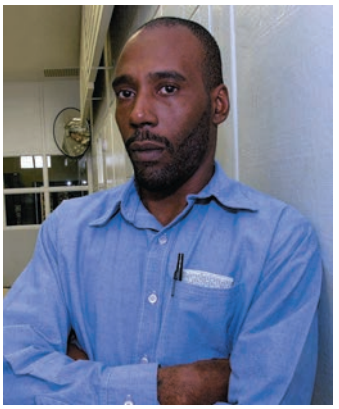
In a murder case, for example, this method can be used to compare DNA samples from the suspect, the victim, and a small amount of blood found at the crime scene. The forensic scientist tests only a few selected portions of the DNA—usually 13 STR markers. However, even this small set of markers can provide a forensically useful genetic profile because the probability that two people (who are not identical twins) would have exactly the same set of STR markers is vanishingly small. The Innocence Project, a nonprofit organization dedicated to overturning wrongful convictions, uses STR analysis of archived samples from crime scenes to revisit old cases. As of 2016, more than 340 innocent people have been released from prison as a result of forensic and legal work by this group **(Figure 19.24)**.

Genetic profiles can also be useful for other purposes. A comparison of the DNA of a mother, her child, and the purported father can conclusively settle a question of paternity. Sometimes paternity is of historical interest: Genetic profiles provided strong evidence that Thomas Jefferson or one of his close male relatives fathered at least one of the children of his slave Sally Hemings. Genetic profiles can also identify victims of mass casualties. The largest such effort occurred after the attack on the World Trade Center in 2001; more than 10,000 samples of victims' remains were compared with DNA samples from personal items, such as toothbrushes, provided by families. Ultimately, forensic scientists succeeded in identifying almost 3,000 victims using these methods.

Just how reliable is a genetic profile? The greater the number of markers examined in a DNA sample, the more likely it is that the profile is unique to one individual. In forensic cases using STR analysis with 13 markers, the probability of two people having identical DNA profiles is somewhere between one chance in 10 billion and one in several trillion. (For comparison, the world's population is between 7 and 8 billion.) The exact probability depends on the frequency of those markers in the general population. Information on how common various markers are in different ethnic groups is critical because these marker frequencies may vary considerably among ethnic groups and between a particular ethnic group and the population as a whole. With the increasing availability of frequency data, forensic scientists can make extremely accurate statistical calculations. Thus, despite problems that can still arise from insufficient data, human error, or flawed evidence, genetic profiles are now accepted as compelling evidence by legal experts and scientists alike.

Figure 19.24 STR analysis used to release an innocent man from prison.

(a) In 1984, Earl Washington was convicted and sentenced to death for the 1982 rape and murder of Rebecca Williams. His sentence was commuted to life in prison in 1993 due to new doubts about the evidence. In 2000, STR analysis by forensic scientists associated with the Innocence Project showed conclusively that he was innocent. This photo shows Washington just before his release in 2001, after 17 years in prison.



Source of sample	STR marker 1	STR marker 2	STR marker 3
Semen on victim	17,19	13,16	12,12
Earl Washington	16,18	14,15	11,12
Kenneth Tinsley	17,19	13,16	12,12

(b) In STR analysis, selected STR markers in a DNA sample are amplified by PCR, and the PCR products are separated by electrophoresis. The procedure reveals how many repeats are present for each STR locus in the sample. An individual has two alleles per STR locus, each with a certain number of repeats. This table shows the number of repeats for three STR markers in three samples: from semen found on the victim, from Washington, and from another man (Kenneth Tinsley), who was in prison because of an unrelated conviction. These and other STR data (not shown) exonerated Washington and led Tinsley to plead guilty to the murder.

Environmental Cleanup

Increasingly, the diverse abilities of certain microorganisms to transform chemicals are being exploited for environmental cleanup. If the growth needs of such microorganisms make them unsuitable for direct use, scientists can now transfer the genes for their valuable metabolic capabilities into other microorganisms, which can then be used to treat environmental problems. For example, many bacteria can extract heavy metals, such as copper, lead, and nickel, from their environments and incorporate the metals into compounds such as copper sulfate or lead sulfate, which are readily recoverable. Genetically engineered microorganisms may become important in both mining (especially as ore reserves are depleted) and cleaning up highly toxic mining wastes. Biotechnologists are also trying to engineer microorganisms that can degrade chlorinated hydrocarbons and other harmful compounds.

These microorganisms could be used in wastewater treatment plants or by manufacturers before the compounds are ever released into the environment.

Agricultural Applications

Scientists are working to learn more about the genomes of agriculturally important plants and animals. For a number of years, they have been using DNA technology in an effort to improve agricultural productivity. The selective breeding of both livestock (animal husbandry) and crops has exploited naturally occurring mutations and genetic recombination for thousands of years.

As we described earlier, DNA technology enables scientists to produce transgenic animals, which speeds up the selective breeding process. The goals of creating a transgenic animal are often the same as the goals of traditional breeding—for instance, to make a sheep with better quality wool, a pig with leaner meat, or a cow that will mature in a shorter time. Scientists might, for example, identify and clone a gene that causes the development of larger muscles (muscles make up most of the meat we eat) in one breed of cattle and transfer it to other cattle or even to sheep. However, health problems are not uncommon among farm animals carrying genes from other species, and modification of the animal's own genes using the CRISPR-Cas9 system will likely emerge as a more useful technique. Animal health and welfare are important issues to consider when genetically altering animals.

Agricultural scientists have already endowed a number of crop plants with genes for desirable traits, such as delayed ripening and resistance to spoilage, disease, and drought. Modifications can also add value to food crops, giving them a longer shelf life or improved flavor or nutritional value. For many plant species, a single tissue cell grown in culture can give rise to an adult plant. Thus, genetic manipulations can be performed on an ordinary somatic cell and the cell then used to generate a plant with new traits.

Genetic engineering is rapidly replacing traditional plant-breeding programs, especially for useful traits, such as herbicide or pest resistance, determined by one or a few genes. Crops engineered with a bacterial gene making the plants resistant to an herbicide can grow while weeds are destroyed, and genetically engineered crops that can resist destructive insects reduce the need for chemical insecticides. In India, the insertion of a salinity resistance gene from a coastal mangrove plant into the genomes of several rice varieties has resulted in rice plants that can grow in water three times as salty as seawater. The research foundation that carried out this feat of genetic engineering estimates that one-third of all irrigated land has high salinity owing to overirrigation and intensive use of chemical fertilizers, representing a serious threat to the food supply. Thus, salinity-resistant crop plants would be enormously valuable worldwide.

Safety and Ethical Questions Raised by DNA Technology

Early concerns about potential dangers associated with recombinant DNA technology focused on the possibility that hazardous new pathogens might be created. What might happen, for instance, if in a research study cancer cell genes were transferred into bacteria or viruses? To guard against such rogue microorganisms, scientists developed a set of guidelines that were adopted as formal government regulations in the United States and some other countries. One safety measure is a set of strict laboratory procedures designed to prevent engineered microorganisms from infecting researchers or accidentally leaving the laboratory. In addition, strains of microorganisms to be used in recombinant DNA experiments are genetically crippled to ensure that they cannot survive outside the laboratory. Finally, certain obviously dangerous experiments have been banned.

Today, most public concern about possible hazards centers not on recombinant microorganisms but on **genetically modified organisms (GMOs)** used as food. A GMO is a transgenic organism, one that has acquired by artificial means one or more genes from another species or even from another variety of the same species. Some salmon, for example, have been genetically modified by addition of a more active salmon growth hormone gene. However, the majority of the GMOs that contribute to our food supply are not animals, but crop plants.

GM crops are widespread in the United States, Argentina, and Brazil; together, these countries account for over 80% of the world's acreage devoted to such crops. In the United States, most corn, soybean, and canola crops are genetically modified, and a recent law requires labeling of GM products. The same foods are an ongoing subject of controversy in Europe, where the GM revolution has been met with strong opposition. Many Europeans are concerned about the safety of GM foods and the possible environmental consequences of growing GM plants. Although a small number of GM crops have been grown on European soil, the European Union established a comprehensive legal framework regarding GMOs in 2015. Among other regulations, individual member states may ban either the growing or importing of GM crops, which must be clearly labeled. The high degree of consumer distrust in Europe makes the future of GM crops there uncertain.

Advocates of a cautious approach toward GM crops fear that transgenic plants might pass their new genes to close relatives in nearby wild areas. We know that lawn and crop grasses, for example, commonly exchange genes with wild relatives via pollen transfer. If crop plants carrying genes for resistance to herbicides, diseases, or insect pests pollinated wild ones, the offspring might become "super weeds" that are very difficult to control. Another worry involves possible risks to human health from GM foods. Some people fear that the protein products of transgenes might lead to allergic

reactions. Although there is some evidence that this could happen, advocates claim that these proteins could be tested in advance to avoid producing ones that cause allergic reactions. (For further discussion of plant biotechnology and GM crops, see Concept 38.3.)

Today, governments and regulatory agencies throughout the world are grappling with how to facilitate the use of biotechnology in agriculture, industry, and medicine while ensuring that new products and procedures are safe. In the United States, such applications of biotechnology must be evaluated for potential risks by various regulatory agencies, including the Food and Drug Administration, the Environmental Protection Agency, the National Institutes of Health, and the Department of Agriculture. Meanwhile, these same agencies and the public must consider the ethical implications of biotechnology.

ABC News Video: Genetically Altered Salmon

Advances in biotechnology have allowed us to obtain complete genome sequences for humans and many other species, providing a vast treasure trove of information about genes. We can ask how certain genes differ from species to species, as well as how genes and, ultimately, entire genomes have evolved. (These are the subjects of Chapter 20.) At the same time, the increasing speed and falling cost of sequencing the genomes of individuals are raising significant ethical questions. Who should have the right to examine someone else's genetic information? How should that information be used? Should a person's genome be a factor in determining eligibility for a job or insurance? Ethical considerations, as well as concerns about potential environmental and health hazards, will likely slow some applications of biotechnology. There is always a danger that too much regulation will stifle basic research and its potential benefits. On the other hand, genetic engineering—especially gene editing with the CRISPR-Cas system—enables us to profoundly and rapidly alter species that have been evolving for millennia. A good example is the potential use of a gene drive that would eliminate the ability of mosquito species to carry diseases or even eradicate certain mosquito species. There would probably be health benefits to this approach, at least initially, but unforeseen problems could easily arise. Given the tremendous power of DNA technology, we must proceed with humility and caution.

Interview with David Suzuki: Exploring ethical issues related to DNA technology

CONCEPT CHECK 19.4

1. What is the advantage of using stem cells for gene therapy or gene editing?
2. List at least three different properties that have been acquired by crop plants via genetic engineering.
3. **WHAT IF?** ➤ As the investigator of a murder case, how can you use DNA technology to identify the guilty? Discuss the genetic basis of this technology.

For suggested answers, see Appendix A.

19 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

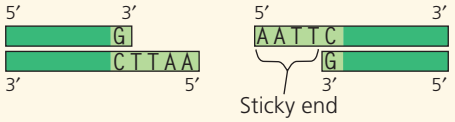
SUMMARY OF KEY CONCEPTS

CONCEPT 19.1

DNA sequencing and DNA cloning are valuable tools for genetic engineering and biological inquiry *(pp. 448–455)*

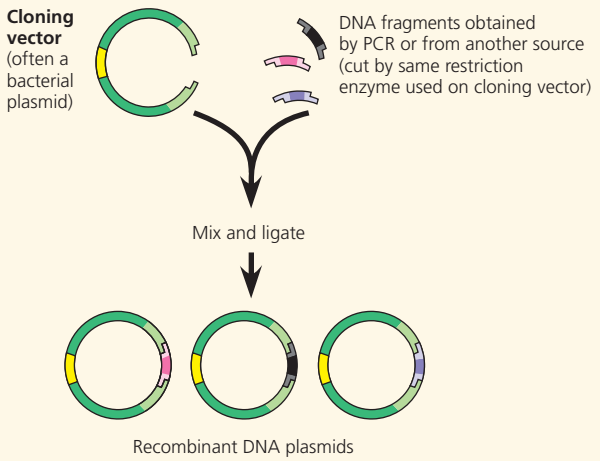
VOCAB SELF-QUIZ goo.gl/Rn5Uax

- **Nucleic acid hybridization**, the base pairing of one strand of a nucleic acid to the complementary sequence on a strand from another nucleic acid molecule, is widely used in **DNA technology**.
- **DNA sequencing** can be carried out using the dideoxy chain termination method in automated sequencing machines.
- Next-generation (high-throughput) techniques for sequencing DNA are based on sequencing by synthesis: DNA polymerase is used to synthesize a stretch of DNA from a single-stranded template, and the order in which nucleotides are added reveals the sequence. Third-generation sequencing methods, including nanopore technology, sequence long DNA molecules one at a time by distinguishing the nucleotide bases as they pass through a pore in a membrane.
- **Gene cloning** (or **DNA cloning**) produces multiple copies of a gene (or DNA segment) that can be used to manipulate and analyze DNA and to produce useful new products or organisms with beneficial traits.
- In **genetic engineering**, bacterial **restriction enzymes** are used to cut DNA molecules within short, specific nucleotide sequences (**restriction sites**), yielding a set of double-stranded **restriction fragments** with single-stranded **sticky ends**:



- The sticky ends on restriction fragments from one DNA source can base-pair with complementary sticky ends on fragments from other DNA molecules. Sealing the base-paired fragments with DNA ligase produces **recombinant DNA molecules**.
- DNA restriction fragments of different lengths can be separated by **gel electrophoresis**.

- The **polymerase chain reaction (PCR)** can produce many copies of (amplify) a specific target segment of DNA *in vitro*, using primers that bracket the desired sequence and a heat-resistant DNA polymerase.
- To clone a eukaryotic gene:



  Recombinant plasmids are returned to host cells, each of which divides to form a clone of cells.
- Several technical difficulties hinder the expression of cloned eukaryotic genes in bacterial host cells. The use of cultured eukaryotic cells as host cells, coupled with appropriate **expression vectors**, helps avoid these problems.

? *Describe how the process of gene cloning results in a cell clone containing a recombinant plasmid.*

CONCEPT 19.2

Biologists use DNA technology to study gene expression and function *(pp. 455–460)*

- Several techniques use hybridization of a **nucleic acid probe** to detect the presence of specific mRNAs.
- ***In situ* hybridization** and **RT-PCR** can detect the presence of a given mRNA in a tissue or an RNA sample, respectively.
- **DNA microarrays** are used to identify sets of genes co-expressed by a group of cells. Increasingly, instead, **RNA sequencing (RNA-seq)** is used to sequence the **cDNAs** corresponding to RNAs from the cells.
- For a gene of unknown function, experimental inactivation of the gene (a gene knockout) and observation of the resulting phenotypic effects can provide clues to its function. The **CRISPR-Cas9 system** allows researchers to edit genes in living cells in a specific, desired way. The new alleles can be altered so that they are inherited in a biased way through a population (**gene drive**). In humans, **genome-wide association studies** identify and use **single nucleotide polymorphisms (SNPs)** as genetic markers for alleles that are associated with particular conditions.

? *What useful information is obtained by detecting expression of specific genes?*

CONCEPT 19.3

Cloned organisms and stem cells are useful for basic research and other applications *(pp. 460–465)*

- The question of whether all the cells in an organism have the same genome prompted the first attempts at organismal cloning.
- Single differentiated cells from plants are often **totipotent**: capable of generating all the tissues of a complete new plant.
- Transplantation of the nucleus from a differentiated animal cell into an enucleated egg can sometimes give rise to a new animal.
- Certain embryonic **stem cells** (ES cells) from animal embryos and particular adult stem cells from adult tissues can reproduce and differentiate both in the lab and in the organism, offering the potential for medical use. ES cells are **pluripotent** but difficult to acquire. Induced pluripotent stem (iPS) cells resemble ES cells in their capacity to differentiate; they can be generated by reprogramming differentiated cells. iPS cells hold promise for medical research and regenerative medicine.

? *Describe how, using mice, a researcher could carry out (1) organismal cloning, (2) production of ES cells, and (3) generation of iPS cells, focusing on how the cells are reprogrammed. (The procedures are basically the same in humans and mice.)*

CONCEPT 19.4

The practical applications of DNA-based biotechnology affect our lives in many ways *(pp. 465–471)*

- DNA technology, including the analysis of genetic markers such as SNPs, is increasingly being used in the diagnosis of genetic and other diseases and offers potential for better treatment of genetic disorders or even permanent cures through **gene therapy**, or gene editing with the CRISPR-Cas9 system. It also enables more informed cancer therapies. DNA technology is used with cell cultures in the large-scale production of protein hormones and other proteins with therapeutic uses. Some therapeutic proteins are being produced in **transgenic** "pharm" animals.
- Analysis of genetic markers such as **short tandem repeats (STRs)** in DNA isolated from tissue or body fluids found at crime scenes leads to a **genetic profile**. Use of genetic profiles can provide definitive evidence that a suspect is innocent or strong evidence of guilt. Such analysis is also useful in parenthood disputes and in identifying the remains of crime victims.
- Genetically engineered microorganisms can be used to extract minerals from the environment or degrade various types of toxic waste materials.
- The aims of developing transgenic plants and animals are to improve agricultural productivity and food quality.
- The potential benefits of genetic engineering must be carefully weighed against the potential for harm to humans or the environment.

? *What factors affect whether a given genetic disease would be a good candidate for successful gene therapy?*

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–8 can be found in the Study Area in MasteringBiology.

9. **MAKE CONNECTIONS** Imagine you want to study one of the human crystallins, proteins present in the lens of the eye (see Figure 1.8). To obtain a sufficient amount of the protein of interest, you decide to clone the gene that codes for it. Assume you know the sequence of this gene. Explain how you would go about this.

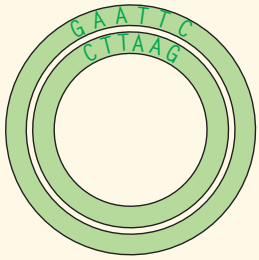
PRACTICE TEST goo.gl/iAsVgL

10. **MAKE CONNECTIONS** Looking at Figure 19.15, what does it mean for a SNP to be "linked" to a disease-associated allele? How does this allow the SNP to be used as a genetic marker? (See Concept 15.3.)

11. DRAW IT You are cloning an aardvark gene, using a bacterial plasmid as a vector. The green diagram shows the plasmid, which contains the restriction site for the enzyme used in Figure 19.5. Above the plasmid is a segment of linear aardvark DNA that was synthesized using PCR. Diagram your cloning procedure, and show what would happen to these two molecules during each step. Use one color for the aardvark DNA and its bases and another color for those of the plasmid. Label each step and all 5′ and 3′ ends.

5′ GAATTCTAAAGCGCTTATGAATTC 3′
3′ CTTAAGATTTCGCGAATACTTAAG 5′

Aardvark DNA



Plasmid

12. EVOLUTION CONNECTION Ethical considerations aside, if DNA-based technologies became widely used, discuss how they might change the way evolution proceeds, as compared with the natural evolutionary mechanisms that have operated for the past 4 billion years.

13. SCIENTIFIC INQUIRY In an experiment aimed at cloning pets, a group of scientists used nuclear transplantation. However, most of the progeny obtained from the experiment were excessively obese, carried many birth defects, and had premature deaths. After a thorough check on the protocols followed, the scientists found some lacunae. What could have gone wrong with the experiment?

14. WRITE ABOUT A THEME: INFORMATION In a short essay (100–150 words), discuss how the genetic basis of life plays a central role in biotechnology.

15. SYNTHESIZE YOUR KNOWLEDGE

The water in the Yellowstone National Park hot springs shown here is around 160°F (70°C). Biologists assumed that no species of organisms could live in water above about 130°F (55°C), so they were surprised to find several species of bacteria there, now called *thermophiles* ("heat-lovers"). You've learned in this chapter how an enzyme from one species, *Thermus aquaticus*, made feasible one of the most important DNA-based techniques used in labs today. Identify the enzyme, and indicate the value of its being isolated from a thermophile. Suggest other reasons why enzymes from this bacterium (or other thermophiles) might also be valuable.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

The Evolution of Genomes

20

Figure 20.1 What genomic information distinguishes a human from a chimpanzee?

KEY CONCEPTS

Elephant shark (*Callorhinchus milii*)

Reading the Leaves from the Tree of Life

The chimpanzee (*Pan troglodytes*) is our closest living relative on the evolutionary tree of life. The boy in **Figure 20.1** and his chimpanzee companion are intently studying the same leaf, but only one of them is able to talk about it. What accounts for this difference between two primates that share so much of their evolutionary history? With advances in sequencing technology, we are now addressing the genetic basis of such intriguing questions. Later in the chapter, you'll learn about the *FOXP2* gene involved in vocalization, which differs between the two species.

The chimpanzee genome was sequenced two years after sequencing of the human genome. Now that we can compare our genome, base by base, with that of the chimpanzee, we can tackle the more general issue of what differences in genetic information account for the distinct characteristics of these two species of primates.

In addition to determining the sequences of the human and chimpanzee genomes, researchers have obtained complete genome sequences for *Escherichia coli* (*E. coli*) and numerous other prokaryotes, as well as many eukaryotes, including *Zea mays* (corn), *Drosophila melanogaster* (fruit fly), *Octopus bimaculoides* (California two-spot octopus), and *Callorhinchus milii* (elephant shark; see the small photo). In 2014, a high-quality sequence was announced for the genome of *Homo neanderthalensis,* (Neanderthals), an extinct species closely related to present-day humans. These genomes are of great interest in their own right, but they also provide important insights into evolution as well as other biological processes. Broadening the

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

human-chimpanzee comparison to the genomes of other primates and more distantly related animals should reveal the sets of genes that control group-defining characteristics. Beyond that, comparisons with the genomes of bacteria, archaea, fungi, protists, and plants should enlighten us about the long evolutionary history of the ancient genes we all share.

With the genomes of many species fully sequenced, scientists can study whole sets of genes and their interactions, an approach called **genomics**. The sequencing efforts that feed this approach have generated, and continue to generate, enormous volumes of data. The need to deal with this ever-increasing flood of information has spawned the field of **bioinformatics**, the application of computational methods to store and analyze biological data.

We will begin this chapter by discussing two approaches to genome sequencing and some of the advances in bioinformatics and its applications. We will then summarize what has been learned from the genomes that have been sequenced thus far. Next, we will describe the composition of the human genome as a representative genome of a complex multicellular eukaryote. Finally, we will explore current ideas about how genomes evolve and about how the evolution of developmental mechanisms could have generated the great diversity of life on Earth today.

CONCEPT 20.1

The Human Genome Project fostered development of faster, less expensive sequencing techniques

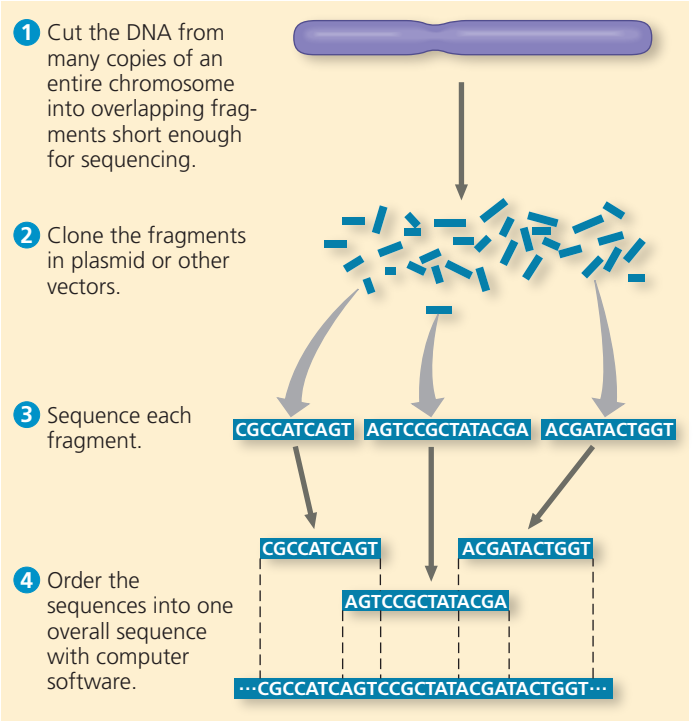
Sequencing of the human genome, an ambitious undertaking, officially began as the **Human Genome Project** in 1990. Organized by an international, publicly funded consortium of scientists at universities and research institutes, the project involved 20 large sequencing centers in six countries plus a host of other labs working on smaller parts of the project.

After the human genome sequence was largely completed in 2003, the sequence of each chromosome was analyzed and described in a series of papers, the last of which covered chromosome 1 and was published in 2006. At that point, the sequencing was considered "virtually complete."

The ultimate goal in mapping any genome is to determine the complete nucleotide sequence of each chromosome. For the human genome, this was accomplished by sequencing machines using the dideoxy chain termination method mentioned in Concept 19.1. Even with this initial use of automation, though, the sequencing of all 3 billion base pairs in a haploid set of human chromosomes presented a formidable challenge. In fact, a major thrust of the Human Genome Project was the development of technology for faster sequencing (see Concept 19.1). Improvements over the years chipped away at each time-consuming step, enabling the rate of sequencing to accelerate impressively: Whereas a productive lab could typically sequence 1,000 base pairs a day in the 1980s, by the year 2000 each research center working on the Human Genome Project was sequencing 1,000 base pairs *per second*. As of 2016, the most widely used automated machines can sequence nearly 25 million base pairs per second, while developers of some newer techniques claim they can achieve a rate of 66 billion base pairs per second. Methods that can analyze biological materials very rapidly and produce enormous volumes of data are said to be "high-throughput." Sequencing machines are an example of high-throughput devices.

Two approaches complemented each other in obtaining the complete sequence. The initial approach was a methodical one that built on an earlier storehouse of human genetic information. In 1998, however, molecular biologist J. Craig Venter set up a company (Celera Genomics) and declared his intention to sequence the entire human genome using an alternative strategy. The **whole-genome shotgun approach** starts with the cloning and sequencing of DNA fragments from randomly cut DNA. Powerful computer programs then assemble the resulting very large number of overlapping short sequences into a single continuous sequence **(Figure 20.2)**.

Figure 20.2 Whole-genome shotgun approach to sequencing. In this approach, developed by J. Craig Venter and colleagues at Celera Genomics, random DNA fragments are cloned (see Figure 19.4), sequenced, and then ordered relative to each other.



VISUAL SKILLS *The fragments in step 2 of this figure are depicted as scattered, rather than arranged in an ordered array. How does this depiction reflect the approach?*

MB **HHMI Animation: Shotgun Sequencing** hhmi BioInteractive

Today, the whole-genome shotgun approach is still used, although newer, "next-generation" sequencing techniques (see Figure 19.3) have resulted in massive increases in speed and decreases in the cost of sequencing entire genomes. In these new techniques, many very small DNA fragments (each about 300 base pairs long) are sequenced at the same time, and computer software rapidly assembles the complete sequence. Because of the sensitivity of these techniques, the fragments can be sequenced directly; the cloning step (2 in Figure 20.2) is unnecessary. Whereas sequencing the first human genome took 13 years and cost $100 million, the genome of James Watson (co-discoverer of DNA structure) was sequenced using newer techniques in four months in 2007 for about $1 million, and as of 2016, an individual's genome can be sequenced in a day or so for about $1,000.

These technological advances have also facilitated an approach called **metagenomics** (from the Greek *meta*, beyond), in which DNA from an entire community of species (a *metagenome*) is collected from an environmental sample and sequenced. Again, computer software sorts out the partial sequences and assembles them into the individual specific genomes. An advantage of this technique is the ability to sequence the DNA of mixed microbial populations, which eliminates the need to culture each species separately in the lab, a difficulty that has limited the study of microbes. So far, this approach has been applied to communities found in environments as diverse as the human intestine and ancient soils in the Arctic, where a 2014 study characterized dozens of species living together as a community as long as 50,000 years ago, including animals and plants as well as microbes.

At first glance, genome sequences of humans and other organisms are simply dry lists of nucleotide bases—millions of A's, T's, C's, and G's in mind-numbing succession. Making sense of this massive amount of data has called for new analytical approaches, which we discuss next.

CONCEPT CHECK 20.1

1. Describe the whole-genome shotgun approach.

For suggested answers, see Appendix A.

CONCEPT 20.2

Scientists use bioinformatics to analyze genomes and their functions

Each of the 20 or so sequencing centers around the world working on the Human Genome Project churned out voluminous amounts of DNA sequence day after day. As the data began to accumulate, the need to coordinate efforts to keep track of all the sequences became clear. Thanks to the foresight of research scientists and government officials involved in the Human Genome Project, its goals included establishing centralized databases and refining analytical software, all to be made readily accessible on the Internet.

HHMI Video: Leading Edge Bioinformatics

Centralized Resources for Analyzing Genome Sequences

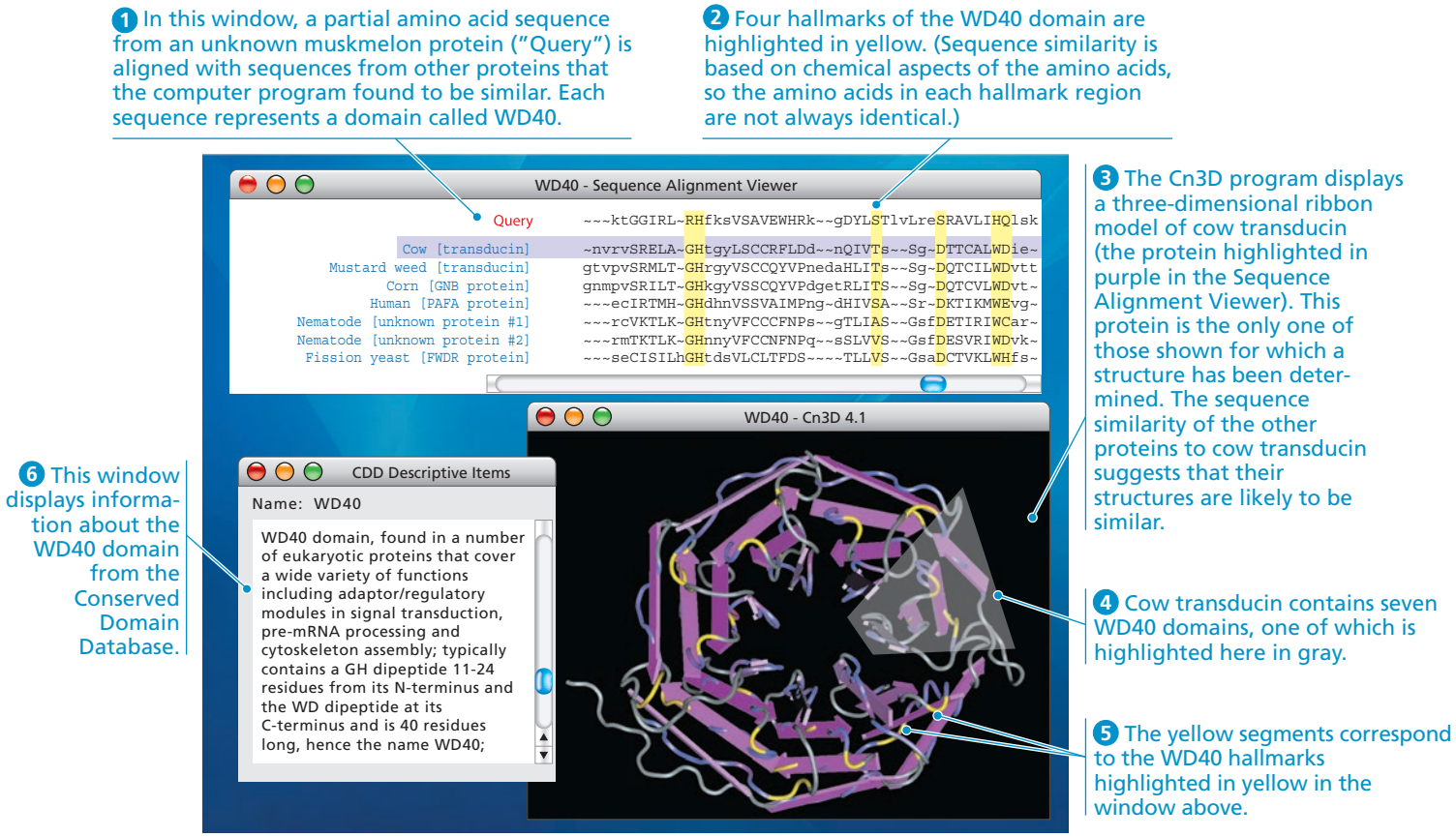
Making bioinformatics resources available to researchers worldwide and speeding up the dissemination of information served to accelerate progress in DNA sequence analysis. For example, in 1988, in preparation for the Human Genome Project in the United States, the National Library of Medicine (NLM) and the National Institutes of Health (NIH) joined forces to create an organization called the National Center for Biotechnology Information (NCBI), which today maintains a website (www.ncbi.nlm.nih.gov) with extensive resources useful for bioinformatics. On this site are links to databases, software, and a wealth of information about genomics and related topics. Similar websites have also been established by three genome centers with which the NCBI collaborates: the European Molecular Biology Laboratory, the DNA Data Bank of Japan, and BGI (formerly known as the Beijing Genomics Institute) in Shenzhen, China. These large, comprehensive websites are complemented by others maintained by individual or small groups of laboratories. Smaller websites often provide databases and software designed for a narrower purpose, such as studying genetic and genomic changes in one particular type of cancer.

The NCBI database of sequences is called GenBank. As of June 2016, it included the sequences of 194 million fragments of genomic DNA, totaling 213 billion base pairs! GenBank is constantly updated, and the amount of data it contains increases rapidly. Any sequence in the database can be retrieved and analyzed using software from the NCBI website or elsewhere.

One very widely used software program available on the NCBI website, called BLAST, allows the user to compare a DNA sequence with every sequence in GenBank, base by base. A researcher might search for similar regions in other genes of the same species or among the genes of other species. Another program allows comparison of protein sequences. Yet a third can search any protein sequence for *conserved* (common) stretches of amino acids (domains) for which a function is known or suspected, and it can show a three-dimensional model of the domain alongside other relevant information **(Figure 20.3)**. There is even a software program that can align and compare a collection of sequences, either nucleic acids or polypeptides, and diagram them in the form of an evolutionary tree based on the sequence relationships. (One such diagram is shown in Figure 20.17.)

Two research institutions, Rutgers University and the University of California, San Diego, also maintain a worldwide database of all three-dimensional protein structures that

Figure 20.3 Bioinformatics tools that are available on the Internet. A website maintained by the National Center for Biotechnology Information (NCBI) allows scientists and the public to access DNA and protein sequences and other stored data. The site includes a link to a protein structure database (Conserved Domain Database, CDD) that can find and describe similar domains in related proteins, as well as software (Cn3D, "See in 3-D") that displays models of domains. Some results are shown from a search for regions of proteins similar to an amino acid sequence in a muskmelon protein. The WD40 domain is very common in proteins encoded by eukaryotic genomes. It often plays a key role in molecular interactions during signal transduction.



have been experimentally determined, called the Protein Data Bank (www.wwpdb.org). These structures can be rotated by the viewer to show all sides of the protein. Throughout this book, you'll find images of protein structures that have been obtained from the Protein Data Bank.

There is a vast array of resources available for researchers anywhere in the world to use free of charge. Let us now consider the types of questions scientists can address using these resources.

Instructors: BLAST Data Analysis Tutorials, which teach students how to work with real data from the BLAST database, can be assigned in MasteringBiology.

Identifying Protein-Coding Genes and Understanding Their Functions

Using available DNA sequences, geneticists can study genes directly, rather than taking the classical genetic approach, which requires determining the function of an unknown gene from the phenotype. But this more recent approach poses a new challenge: What does the gene actually do? Given a long DNA sequence from a database like GenBank, scientists aim to identify all protein-coding genes in the sequence and ultimately their functions. This process, called **gene annotation**, uses three lines of evidence to identify a gene.

First, computers are utilized in a search for patterns that indicate the presence of genes. The usual approach is to use software to scan the stored sequences for those that represent transcriptional and translational start and stop signals, RNA-splicing sites, and other telltale signs of protein-coding genes. The software also looks for certain short sequences that specify known mRNAs. Thousands of such sequences, called *expressed sequence tags*, or *ESTs*, have been collected from cDNA sequences and are cataloged in computer databases. This type of analysis identifies sequences that may turn out to be previously unknown protein-coding genes.

Although the identities of about half of the human genes were known before the Human Genome Project began, the other genes, previously unknown, were revealed by DNA sequence analysis. Once such suspected genes are identified, the second step is to obtain clues about their identities and functions by using software to compare their sequences with those of known genes from other organisms. Due to

redundancy in the genetic code, the DNA sequence itself may vary more among species than the protein sequence does. Thus, scientists interested in proteins often compare the predicted amino acid sequence of a protein to that of other proteins. Third, the identities of these genes must then be confirmed by using RNA-seq (see Figure 19.13) or some other method to show that the relevant RNA is actually expressed from the proposed gene.

Animation: The Human Genome Project: Genes on Human Chromosome 17

Sometimes a newly identified sequence will match, at least partially, the sequence of a gene or protein in another species whose function is well known. For example, a plant researcher working on signaling pathways in the muskmelon would be excited to see that a partial amino acid sequence from a gene she had identified matched sequences in other species encoding a functional part of a protein called a WD40 domain (see Figure 20.3). WD40 domains are present in many eukaryotic proteins and are known to function in signal transduction pathways. Alternatively, a new gene sequence might be similar to a previously encountered sequence whose function is still unknown. Another possibility is that the sequence is entirely unlike anything ever seen before. This was true for about a third of the genes of *E. coli* when its genome was sequenced. In such cases, protein function is usually deduced through a combination of biochemical and functional studies. The biochemical approach aims to determine the three-dimensional structure of the protein as well as other attributes, such as potential binding sites for other molecules. Functional studies usually involve *knocking out* (blocking or disabling) the gene in an organism to see how the phenotype is affected. The CRISPR-Cas 9 system, described in Figure 19.14, is an example of an experimental technique used to block gene function.

Understanding Genes and Gene Expression at the Systems Level

The impressive computational power provided by the tools of bioinformatics allows the study of whole sets of genes and their interactions, as well as the comparison of genomes from different species. Genomics is a rich source of new insights into fundamental questions about genome organization, regulation of gene expression, embryonic development, and evolution.

One informative approach was taken by a long-term research project called ENCODE (Encyclopedia of DNA Elements), which ran from 2003 to 2012. The aim of the project was to learn everything possible about the functionally important elements in the human genome using multiple experimental techniques on different types of cultured cells. Investigators sought to identify protein-coding genes and genes for non-coding RNAs, along with sequences that regulate gene expression, such as enhancers and promoters. In addition, they extensively characterized DNA and histone modifications and chromatin structure—features termed "epigenetic" since they affect gene expression without changing the sequence of nucleotide bases (see Concept 18.3). The second phase of the project, involving more than 440 scientists in 32 research groups, culminated in 2012 with the simultaneous publication of 30 papers describing over 1,600 large data sets. The considerable power of this project is that it provides the opportunity to compare results from specific projects with each other, yielding a much richer picture of the whole genome.

Perhaps the most striking finding is that about 75% of the genome is transcribed at some point in at least one of the cell types studied, even though less than 2% codes for proteins. Furthermore, biochemical functions have been assigned to DNA elements making up at least 80% of the genome. To learn more about the different types of functional elements, parallel projects are analyzing in a similar way the genomes of two model organisms, the soil nematode *Caenorhabditis elegans* and the fruit fly *Drosophila melanogaster*. Because genetic and biochemical experiments using DNA technology can be performed on these species, testing the activities of potentially functional DNA elements in their genomes is expected to illuminate the workings of the human genome.

Because the ENCODE project analyzed cells in culture, its potential for clinical applications was limited. A related project called the Roadmap Epigenomics Project set out to characterize the *epigenome*—the epigenetic features of the genome—of hundreds of human cell types and tissues. The aim was to focus on the epigenomes of stem cells, normal tissues from mature adults, and relevant tissues from individuals with diseases such as cancer and neurodegenerative and autoimmune disorders. In 2015, a series of papers reported on the results from 111 tissues. One of the most useful findings was that the original tissue in which a cancer arose can be identified in cells of a secondary tumor based on characterization of their epigenomes.

Systems Biology

The scientific progress resulting from sequencing genomes and studying large sets of genes has encouraged scientists to attempt similar systematic studies of sets of proteins and their properties (such as their abundance, chemical modifications, and interactions), an approach called **proteomics**. (A **proteome** is the entire set of proteins expressed by a cell or group of cells.) Proteins, not the genes that encode them, carry out most of the activities of the cell. Therefore, we must study when and where proteins are produced in an organism, as well as how they interact in networks, if we are to understand the functioning of cells and organisms.

Genomics and proteomics enable molecular biologists to approach the study of life from an increasingly global perspective. Using the tools we have described, biologists have begun to compile catalogs of genes and proteins— listings of all the "parts" that contribute to the operation of cells,

tissues, and organisms. With such catalogs in hand, researchers have shifted their attention from the individual parts—the genes and proteins—to their functional integration in biological systems. As you may recall, Concept 1.1 discussed this approach, called **systems biology**, which aims to model the dynamic behavior of whole biological systems based on the study of the interactions among the system's parts. Because of the vast amounts of data generated in these types of studies, advances in computer technology and bioinformatics are crucial to studying systems biology.

One important use of the systems biology approach is to define gene and protein interaction networks. To map the protein interaction network in the yeast *Saccharomyces cerevisiae*, for instance, researchers used sophisticated techniques to knock out pairs of genes, one pair at a time, creating doubly mutant cells. They then compared the fitness of each double mutant (based in part on the size of the cell colony it formed) to that predicted from the fitness of each of the two single mutants. The researchers reasoned that if the observed fitness matched the prediction, then the products of the two genes didn't interact with each other, but if the observed fitness was greater or less than predicted, then the gene products interacted in the cell. They then used computer software to build a graphic model by "mapping" the gene products to certain locations in the model, based on the similarity of their interactions. This resulted in the network-like "functional map" of protein interactions shown in **Figure 20.4**. Processing the vast number of protein-protein interactions generated by this experiment and integrating them into the completed map required powerful computers, mathematical tools, and newly developed software.

Application of Systems Biology to Medicine

The Cancer Genome Atlas is another example of systems biology in which many interacting genes and gene products are analyzed together, as a group. This project, under the joint leadership of the National Cancer Institute and the NIH, aims to determine how changes in biological systems lead to cancer. A three-year pilot project that ended in 2010 set out to find all the common mutations in three types of cancer—lung cancer, ovarian cancer, and glioblastoma of the brain—by comparing gene sequences and patterns of gene expression in cancer cells with those in normal cells. Work on glioblastoma confirmed the role of several suspected genes and identified a few previously unknown ones, suggesting possible new targets for therapies. The approach proved so fruitful for these three types of cancer that it has been extended to ten other types, chosen because they are common and often lethal in humans.

As high-throughput techniques become more rapid and less expensive, they are being increasingly applied to the problem of cancer, as seen with the Roadmap Epigenomics Project described earlier. Rather than sequencing only protein-coding genes, sequencing the whole genomes of many tumors of a particular type allows scientists to uncover common chromosomal abnormalities, as well as any other consistent changes in these aberrant genomes.

Figure 20.4 The systems biology approach to protein interactions. This global protein interaction map shows the likely interactions (lines) among about 4,500 gene products (dots) in *Saccharomyces cerevisiae,* the budding yeast. Dots of the same color represent gene products involved in one of the 13 similarly colored cellular functions listed around the map. The white dots represent proteins that haven't been assigned to any color-coded function. The expanded area shows additional details of one map region where the gene products (blue dots) carry out amino acid biosynthesis, uptake, and related functions.

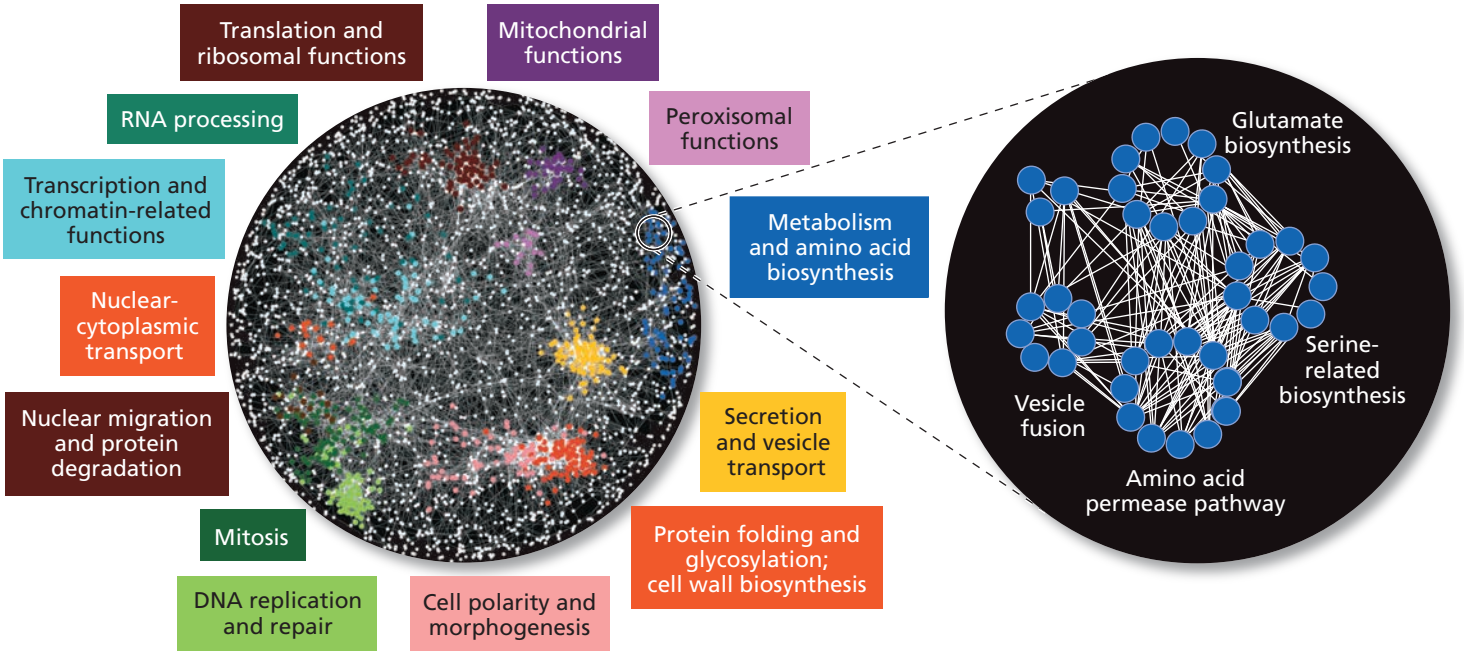


Figure 20.5 A human gene microarray chip. Tiny spots of DNA arranged in a grid on this silicon wafer represent almost all of the genes in the human genome. Using this chip, researchers can analyze expression patterns for all these genes at the same time (see Figure 19.12).

HHMI Animation: Gene Chip Manufacturing

In addition to whole-genome sequencing, silicon and glass "chips" that hold a microarray of most of the known human genes are now used to analyze gene expression patterns in patients suffering from various cancers and other diseases **(Figure 20.5)**. Increasingly, RNA-seq (see Figure 19.13) is replacing microarray analysis. Analyzing which genes are over- or underexpressed in a particular cancer may allow physicians to tailor patients' treatment to their unique genetic makeup and the specifics of their cancers. This approach has been used to characterize subsets of particular cancers, enabling more refined treatments. Breast cancer is one example (see Figure 18.27).

Ultimately, medical records may include an individual's DNA sequence, a sort of genetic bar code, with regions highlighted that predispose the person to specific diseases. The use of such sequences for personalized medicine—disease prevention and treatment—has great potential.

Systems biology is a very efficient way to study emergent properties at the molecular level. Novel properties arise at each successive level of biological complexity as a result of the arrangement of building blocks at the underlying level (see Concept 1.1). The more we can learn about the arrangement and interactions of the components of genetic systems, the deeper will be our understanding of whole organisms. The rest of this chapter surveys what we've learned from genomic studies.

CONCEPT CHECK 20.2

1. What are the ways by which the functions of newly sequenced genes can be ascertained?
2. Explain the advantage of the systems biology approach to studying cancer versus the approach of studying a single gene at a time.
3. **MAKE CONNECTIONS ➤** The ENCODE pilot project found that at least 75% of the genome is transcribed into RNAs, far more than could be accounted for by protein-coding genes. Review Concepts 17.3 and 18.3 and suggest some roles that these RNAs might play.
4. **MAKE CONNECTIONS ➤** In Concept 19.2, you learned about genome-wide association studies. Explain how these studies use the systems biology approach.

For suggested answers, see Appendix A.

CONCEPT 20.3

Genomes vary in size, number of genes, and gene density

The sequences of thousands of genomes have been completed, with tens of thousands of genomes either in progress or considered permanent drafts (because they require more work than it would be worth to complete them). Among the sequences in progress are roughly 3,400 metagenomes. In the completely sequenced group, about 5,000 are genomes of bacteria, and 242 are archaeal genomes. There are 283 completed eukaryotic species, along with 2,635 permanent drafts. Among these are vertebrates, invertebrates, protists, fungi, and plants. Next, we'll discuss what we've learned about genome size, number of genes, and gene density, focusing on general trends.

Genome Size

Comparing the three domains (Bacteria, Archaea, and Eukarya), we find a general difference in genome size between prokaryotes and eukaryotes **(Table 20.1)**. While there are

Table 20.1 Genome Sizes and Estimated Numbers of Genes*

Organism	Haploid Genome Size (Mb)	Number of Genes	Genes per Mb
Bacteria			
Haemophilus influenzae	1.8	1,700	940
Escherichia coli	4.6	4,400	950
Archaea			
Archaeoglobus fulgidus	2.2	2,500	1,130
Methanosarcina barkeri	4.8	3,600	750
Eukaryotes			
Saccharomyces cerevisiae (yeast, a fungus)	12	6,300	525
Urticularia gibba (floating bladderwort)	82	28,500	348
Caenorhabditis elegans (nematode)	100	20,100	200
Arabidopsis thaliana (mustard family plant)	120	27,000	225
Drosophila melanogaster (fruit fly)	165	14,000	85
Daphnia pulex (water flea)	200	31,000	155
Zea mays (corn)	2,300	32,000	14
Ailuropoda melanoleuca (giant panda)	2,400	21,000	9
Homo sapiens (human)	3,000	< 21,000	7
Paris japonica (Japanese canopy plant)	149,000	ND	ND

*Some values given here are likely to be revised as genome analysis continues.
Mb = million base pairs. ND = not determined.

some exceptions, most bacterial genomes have between 1 and 6 million base pairs (Mb); the genome of *E. coli*, for instance, has 4.6 Mb. Genomes of archaea are, for the most part, within the size range of bacterial genomes. (Keep in mind, however, that many fewer archaeal genomes have been completely sequenced, so this picture may change.) Eukaryotic genomes tend to be larger: The genome of the single-celled yeast *Saccharomyces cerevisiae* (a fungus) has about 12 Mb, while most animals and plants, which are multicellular, have genomes of at least 100 Mb. There are 165 Mb in the fruit fly genome, while humans have 3,000 Mb, about 500 to 3,000 times as many as a typical bacterium.

Aside from this general difference between prokaryotes and eukaryotes, a comparison of genome sizes among eukaryotes fails to reveal any systematic relationship between genome size and the organism's phenotype. For instance, the genome of *Paris japonica*, the Japanese canopy plant, contains 149 billion base pairs (149,000 Mb), while that of another plant, *Urticularia gibba*, a bladderwort, contains only 82 Mb. Even more striking, there is a single-celled amoeba, *Polychaos dubium*, whose genome size has been estimated at 670 billion base pairs (670,000 Mb). (This genome has not yet been sequenced.) On a finer scale, comparing two insect species, the cricket (*Anabrus simplex*) genome turns out to have 11 times as many base pairs as the fruit fly (*Drosophila melanogaster*) genome. There is a wide range of genome sizes within the groups of unicellular eukaryotes, insects, amphibians, and plants, and less of a range within mammals and reptiles.

Number of Genes

The number of genes also varies between prokaryotes and eukaryotes: Bacteria and archaea, in general, have fewer genes than eukaryotes. Free-living bacteria and archaea have from 1,500 to 7,500 genes, while the number of genes in eukaryotes ranges from about 5,000 for unicellular fungi (yeasts) to at least 40,000 for some multicellular eukaryotes.

Within the eukaryotes, the number of genes in a species is often lower than expected from considering simply the size of its genome. Looking at Table 20.1, you can see that the genome of the nematode *C. elegans* is 100 Mb in size and contains roughly 20,100 genes. In comparison, the genome of *Drosophila melanogaster* is much bigger (165 Mb) but has only about two-thirds the number of genes—14,000 genes.

Considering an example closer to home, we noted that the human genome contains 3,000 Mb, well over ten times the size of either the *D. melanogaster* or *C. elegans* genome. At the outset of the Human Genome Project, biologists expected somewhere between 50,000 and 100,000 genes to be identified in the completed sequence, based on the number of known human proteins. As the project progressed, the estimate was revised downward several times, and the ENCODE project discussed above has established the number to be fewer than 21,000. This relatively low number, similar to the number of genes in the nematode *C. elegans*, surprised biologists, who had been expecting many more human genes.

Interview with Eric Lander: Exploring the Human Genome

What genetic attributes allow humans (and other vertebrates) to get by with no more genes than nematodes? An important factor is that vertebrate genomes "get more bang for the buck" from their coding sequences because of extensive alternative splicing of RNA transcripts. Recall that this process generates more than one polypeptide from a single gene (see Figure 18.13). A typical human gene contains about ten exons, and an estimated 90% or more of these multi-exon genes are spliced in at least two different ways. Some genes are expressed in hundreds of alternatively spliced forms, others in just two. Scientists have not yet catalogued all of the different forms, but it is clear that the number of different proteins encoded in the human genome far exceeds the proposed number of genes.

Additional polypeptide diversity could result from post-translational modifications such as cleavage or the addition of carbohydrate groups in different cell types or at different developmental stages. Finally, the discovery of miRNAs and other small RNAs that play regulatory roles has added a new variable to the mix (see Concept 18.3). Some scientists think that this added level of regulation, when present, may contribute to greater organismal complexity for some genes.

Gene Density and Noncoding DNA

We can take both genome size and number of genes into account by comparing gene density in different species. In other words, we can ask: How many genes are in a given length of DNA? When we compare the genomes of bacteria, archaea, and eukaryotes, we see that eukaryotes generally have larger genomes but fewer genes in a given number of base pairs. Humans have hundreds or thousands of times as many base pairs in their genome as most bacteria, as we already noted, but only 5 to 15 times as many genes; thus, gene density is lower in humans (see Table 20.1). Even unicellular eukaryotes, such as yeasts, have fewer genes per million base pairs than bacteria and archaea. Among the genomes that have been sequenced completely, humans and other mammals have the lowest gene density.

In all bacterial genomes studied so far, most of the DNA consists of genes for protein, tRNA, or rRNA; the small amount remaining consists mainly of nontranscribed regulatory sequences, such as promoters. The sequence of nucleotides along a bacterial protein-coding gene is not interrupted by noncoding sequences (introns). In eukaryotic genomes, by contrast, most of the DNA neither encodes protein nor is transcribed into RNA molecules of known function, and

the DNA includes more complex regulatory sequences. In fact, humans have 10,000 times as much noncoding DNA as bacteria. Some of this DNA in multicellular eukaryotes is present as introns within genes. Indeed, introns account for most of the difference in average length between human genes (27,000 base pairs) and bacterial genes (1,000 base pairs).

In addition to introns, multicellular eukaryotes have a vast amount of non-protein-coding DNA between genes. In the next section, we will describe the composition and arrangement of these great stretches of DNA in the human genome.

CONCEPT CHECK 20.3

1. Eukaryotes such as humans have around 21,000 genes and 7 genes per megabase. Prokaryotes such as *E. coli*, on the other hand, have 4,400 genes and 950 genes per megabase. What accounts for the lower gene density in eukaryotes?
2. The Genomes Online Database (GOLD) website of the Joint Genome Institute has information about genome sequencing projects. Scroll through the page at https://gold.jgi.doe.gov/statistics and describe the information you find. What percent of bacterial genome projects have medical relevance?
3. **WHAT IF?** ➤ What evolutionary processes might account for prokaryotes having smaller genomes than eukaryotes?

For suggested answers, see Appendix A.

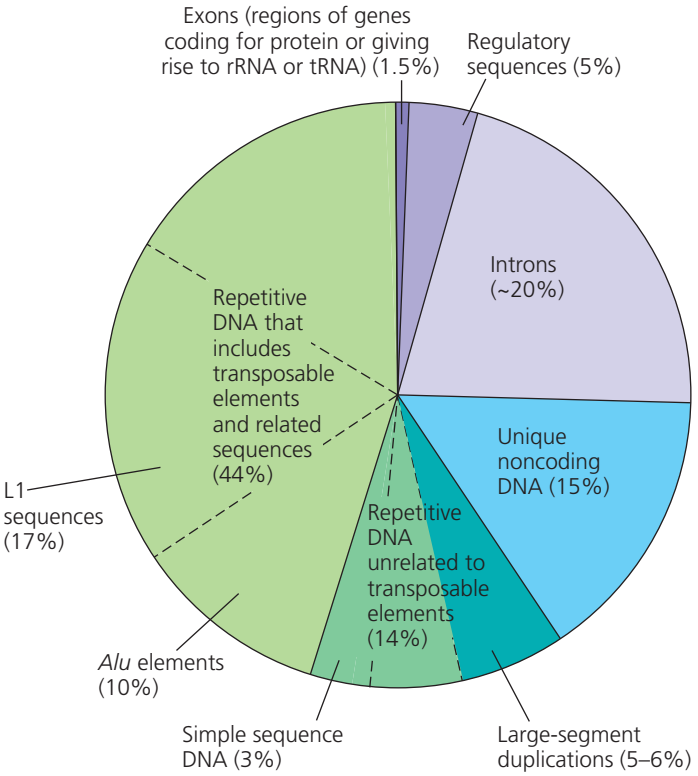
CONCEPT 20.4

Multicellular eukaryotes have a lot of noncoding DNA and many multigene families

We have spent most of this chapter, and indeed this unit, focusing on genes that code for proteins. Yet the coding regions of these genes and the genes for noncoding RNA products such as rRNA, tRNA, and miRNA make up only a small portion of the genomes of most multicellular eukaryotes. For example, once the sequencing of the human genome was completed, it became clear that only a tiny part—about 1.5%—codes for proteins or is transcribed into rRNAs or tRNAs. **Figure 20.6** shows what is known about the makeup of the remaining 98.5% of the genome.

Gene-related regulatory sequences and introns account, respectively, for 5% and about 20% of the human genome. The rest, located between functional genes, includes some unique (single-copy) noncoding DNA, such as gene fragments and **pseudogenes**, former genes that have accumulated mutations over a long time and no longer produce functional proteins. (The genes that produce small noncoding RNAs are a tiny percentage of the genome, distributed between the 20% introns and the 15% unique noncoding DNA.) Most of the DNA between functional genes, however,

Figure 20.6 Types of DNA sequences in the human genome. The gene sequences that code for proteins or are transcribed into rRNA or tRNA molecules make up only about 1.5% of the human genome (dark purple in the pie chart), while introns and regulatory sequences associated with genes (light purple) make up about a quarter. The vast majority of the human genome does not code for proteins (although much of it gives rise to RNAs), and a large amount is repetitive DNA (dark and light green and teal).



is **repetitive DNA**, which consists of sequences that are present in multiple copies in the genome.

The bulk of many eukaryotic genomes consists of DNA sequences that neither code for proteins nor are transcribed to produce RNAs with known functions; this noncoding DNA was often described in the past as "junk DNA." However, genome comparisons over the past 10 years have revealed the persistence of this DNA in diverse genomes over many hundreds of generations. For example, the genomes of humans, rats, and mice contain almost 500 regions of noncoding DNA that are *identical* in sequence in all three species. This is a higher level of sequence conservation than is seen for protein-coding regions in these species, strongly suggesting that the noncoding regions have important functions. The results of the ENCODE project discussed earlier have underscored the key roles played by much of this noncoding DNA. In the next few pages, we examine how genes and noncoding DNA sequences are organized within genomes of multicellular eukaryotes, using the human genome as our main example. Genome organization tells us a lot about how genomes have evolved and continue to evolve, as we'll discuss in Concept 20.5.

Transposable Elements and Related Sequences

Both prokaryotes and eukaryotes have stretches of DNA that can move from one location to another within the genome. These stretches are known as *transposable genetic elements*, or simply **transposable elements**. During the process called *transposition*, a transposable element moves from one site in a cell's DNA to a different target site by a type of recombination process. Transposable elements are sometimes called "jumping genes," but actually they never completely detach from the cell's DNA. Instead, the original and new DNA sites are brought very close together by enzymes and other proteins that bend the DNA. Surprisingly, about 75% of human repetitive DNA (44% of the entire human genome) is made up of transposable elements and sequences related to them.

The first evidence for wandering DNA segments came from American geneticist Barbara McClintock's breeding experiments with Indian corn (maize) in the 1940s and 1950s **(Figure 20.7)**. As she tracked corn plants through multiple generations, McClintock identified changes in the color of corn kernels that made sense only if she postulated the existence of genetic elements capable of moving from other locations in the genome into the genes for kernel color, disrupting the genes so that the kernel color was changed. McClintock's discovery was met with great skepticism and virtually discounted at the time. Her careful work and insightful ideas were finally validated many years later when transposable elements were found in bacteria. In 1983, at the age of 81, McClintock received the Nobel Prize for her pioneering research.

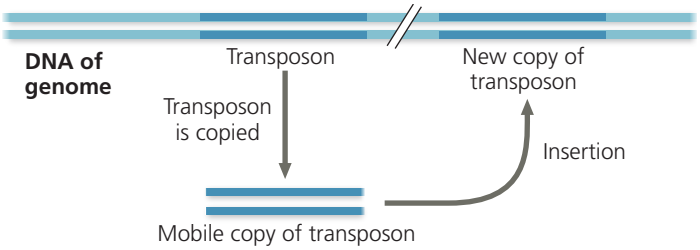
Movement of Transposons and Retrotransposons

Eukaryotic transposable elements are of two types. The first type, **transposons**, move within a genome by means of a DNA intermediate. Transposons can move by a "cut-and-paste" mechanism, which removes the element from the original site, or by a "copy-and-paste" mechanism, which leaves a copy behind **(Figure 20.8)**. Both mechanisms require an enzyme called *transposase*, which is generally encoded by the transposon.

Most transposable elements in eukaryotic genomes are of the second type, **retrotransposons**, which move by means of an RNA intermediate that is a transcript of the retrotransposon DNA. Thus, retrotransposons always leave a copy at the original site during transposition **(Figure 20.9)**. To insert at another site, the RNA intermediate is first converted back to DNA by reverse transcriptase, an enzyme encoded by the retrotransposon. (Reverse transcriptase is also encoded by retroviruses, which are discussed in Concept 26.2. In fact, retroviruses may have evolved from retrotransposons.) Another cellular enzyme catalyzes insertion of the reverse-transcribed DNA at a new site.

Figure 20.7 The effect of transposable elements on corn kernel color. Barbara McClintock first proposed the idea of mobile genetic elements after observing variegations in the color of the kernels on a corn cob (top right).

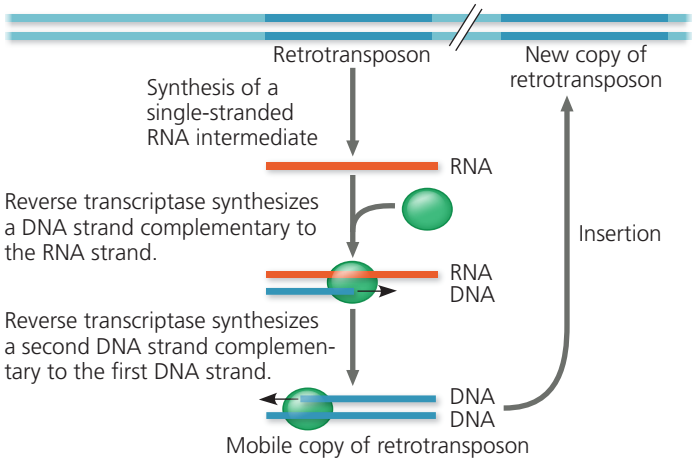
Figure 20.8 Transposon movement. Movement of transposons by either the copy-and-paste mechanism (shown here) or the cut-and-paste mechanism involves a double-stranded DNA intermediate that is inserted into the genome.



VISUAL SKILLS *How would this figure differ if it showed the cut-and-paste mechanism?*

MB **Interview with Virginia Walbot: Plant genetics and development**

Figure 20.9 Retrotransposon movement. Movement begins with synthesis of a single-stranded RNA intermediate. The remaining steps are essentially identical to part of the retrovirus replicative cycle (see Figure 26.8).



Sequences Related to Transposable Elements

Multiple copies of transposable elements and sequences related to them are scattered throughout eukaryotic genomes. A single unit is usually hundreds to thousands of base pairs long, and the dispersed copies are similar but usually not identical to each other. Some of these are transposable elements that can move; the enzymes required for this movement may be encoded by any transposable element, including the one that is moving. Others are related sequences that have lost the ability to move altogether. Transposable elements and related sequences make up 25–50% of most mammalian genomes (see Figure 20.6) and even higher percentages in amphibians and many plants. In fact, the very large size of some plant genomes is accounted for by extra transposable elements rather than by extra genes. For example, transposable elements make up 85% of the corn genome!

In humans and other primates, a large portion of transposable element-related DNA consists of a family of similar sequences called *Alu elements*. These sequences alone account for approximately 10% of the human genome. *Alu* elements are about 300 nucleotides long, much shorter than most functional transposable elements, and they do not code for any protein. However, many *Alu* elements are transcribed into RNA, and at least some of these RNAs are thought to help regulate gene expression.

An even larger percentage (17%) of the human genome is made up of a type of retrotransposon called *LINE-1*, or *L1*. These sequences are much longer than *Alu* elements—about 6,500 base pairs—and typically have a very low rate of transposition. However, researchers working with rats have found L1 retrotransposons to be more active in cells of the developing brain. They have proposed that different effects on gene expression of L1 retrotransposition in developing neurons may contribute to the great diversity of neuronal cell types (see Concept 48.1).

Although many transposable elements encode proteins, these proteins do not carry out normal cellular functions. Therefore, transposable elements are usually included in the "noncoding" DNA category, along with other repetitive sequences.

Other Repetitive DNA, Including Simple Sequence DNA

Repetitive DNA that is not related to transposable elements has probably arisen from mistakes during DNA replication or recombination. Such DNA accounts for about 14% of the human genome (see Figure 20.6). About a third of this (5–6% of the human genome) consists of duplications of long stretches of DNA, with each unit ranging from 10,000 to 300,000 base pairs. These long segments seem to have been copied from one chromosomal location to another site on the same or a different chromosome and probably include some functional genes.

In contrast to scattered copies of long sequences, stretches of DNA known as **simple sequence DNA** contain many copies of tandemly repeated short sequences, as in the following example (showing one DNA strand only):

. . . GTTACGTTACGTTACGTTACGTTACGTTAC . . .

In this case, the repeated unit (GTTAC) consists of 5 nucleotides. Repeated units may contain as many as 500 nucleotides, but often contain fewer than 15 nucleotides, as in this example. When the unit contains 2–5 nucleotides, the series of repeats is called a **short tandem repeat**, or **STR**; we discussed the use of STR analysis in preparing genetic profiles in Concept 19.4 (see Figure 19.24). The number of copies of the repeated unit can vary from site to site within a given genome. There could be as many as several hundred thousand repetitions of the GTTAC unit at one site, but only half that number at another. STR analysis is performed on sites selected because they have relatively few repeats. The repeat number varies from person to person, and since humans are diploid, each person has two alleles per site, which can differ in repeat number. This diversity produces the variation represented in the genetic profiles that result from STR analysis. Altogether, simple sequence DNA makes up 3% of the human genome.

Much of a genome's simple sequence DNA is located at chromosomal telomeres and centromeres, suggesting that this DNA plays a structural role for chromosomes. The DNA at centromeres is essential for the separation of chromatids in cell division (see Concept 12.2). Centromeric DNA, along with simple sequence DNA located elsewhere, may also help organize the chromatin within the interphase nucleus. The simple sequence DNA located at telomeres—the tips of chromosomes—prevents genes from being lost as the DNA shortens with each round of replication (see Concept 16.2). Telomeric DNA also binds proteins that protect the ends of a chromosome from degradation and from joining to other chromosomes.

Short repetitive sequences like those described here provide a challenge for whole-genome shotgun sequencing because the presence of many short repeats hinders accurate reassembly of fragment sequences by computers. Regions of simple sequence DNA account for much of the uncertainty present in estimates of whole-genome sizes and are the reason some sequences are considered "permanent drafts."

Genes and Multigene Families

We finish our discussion of the various types of DNA sequences in eukaryotic genomes with a closer look at genes. Recall that DNA sequences that code for proteins or give rise to tRNA or rRNA compose a mere 1.5% of the human genome (see Figure 20.6). If we include introns and regulatory sequences associated with genes, the total amount of DNA that is gene-related—coding and noncoding—constitutes about 25% of the human genome. Put another way, only about 6% (1.5% out of 25%) of the length of the average gene is represented in the final gene product.

Like the genes of bacteria, many eukaryotic genes are present as unique sequences, with only one copy per haploid set of chromosomes. But in the human genome and the genomes of many other animals and plants, these unique genes make up less than half of the total gene-related DNA. The rest occur in **multigene families**, collections of two or more identical or very similar genes.

In multigene families that consist of *identical* DNA sequences, those sequences are usually clustered tandemly and, with the notable exception of the genes for histone proteins, have RNAs as their final products. An example is the family of identical DNA sequences that each include the genes for the three largest rRNA molecules **(Figure 20.10a)**. These rRNA molecules are transcribed from a single transcription unit that is repeated tandemly hundreds to thousands of times in one or several clusters in the genome of a multicellular eukaryote. The many copies of this rRNA transcription unit help cells to quickly make the millions of ribosomes needed for active protein synthesis. The primary transcript is cleaved to yield three rRNA molecules, which combine with proteins and one other kind of rRNA (5*S* rRNA) to form ribosomal subunits.

The classic examples of multigene families of *nonidentical* genes are two related families of genes that encode globins, a group of proteins that include the α and β polypeptide subunits of hemoglobin. One family, located on chromosome 16 in humans, encodes various forms of α-globin; the other, on chromosome 11, encodes forms of β-globin **(Figure 20.10b)**. The different forms of each globin subunit are expressed at different times in development, allowing hemoglobin to function effectively in the changing environment of the developing animal. In humans, for example, the embryonic and fetal forms of hemoglobin have a higher affinity for oxygen than the adult forms, ensuring the efficient transfer of oxygen from mother to fetus. Also found in the globin gene family clusters are several pseudogenes.

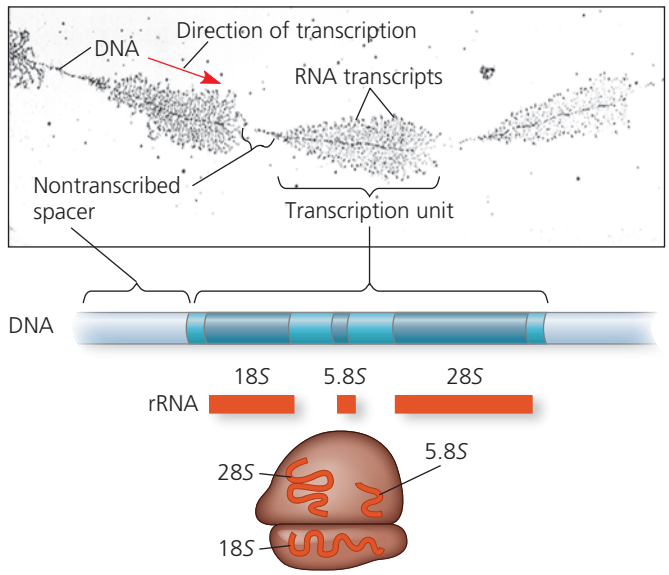
In Concept 20.5, we'll consider the evolution of these two globin gene families as we explore how arrangements of genes provide insight into the evolution of genomes. We'll also examine some processes that have shaped the genomes of different species over evolutionary time.

CONCEPT CHECK 20.4

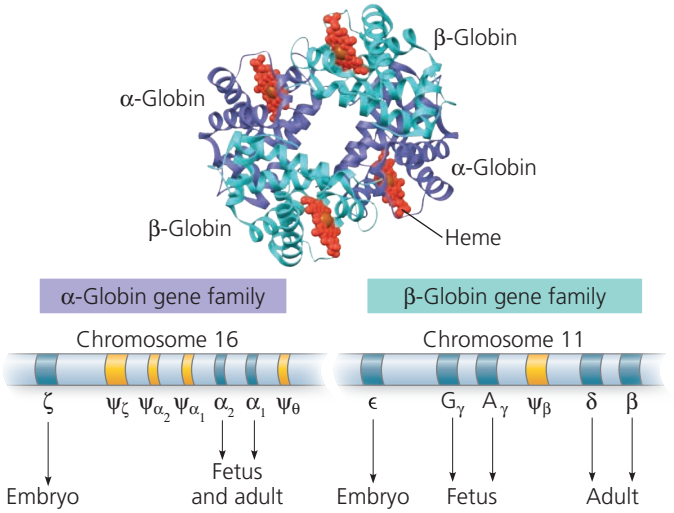
1. What is simple sequence DNA, and where is it located in a chromosome?
2. **VISUAL SKILLS ➤** Which of the three mechanisms described in Figures 20.8 and 20.9 result(s) in a copy remaining at the original site as well as a copy appearing in a new location?
3. Contrast the organizations of the rRNA gene family and the globin gene families. For each, explain how the existence of a family of genes benefits the organism.
4. **MAKE CONNECTIONS ➤** Assign each DNA segment at the top of Figure 18.8 to a sector in the pie chart in Figure 20.6.

For suggested answers, see Appendix A.

Figure 20.10 Gene families.



(a) Part of the ribosomal RNA gene family. The TEM at the top shows three of the hundreds of copies of rRNA transcription units in the rRNA gene family of a salamander genome. Each "feather" corresponds to a single unit being transcribed by about 100 molecules of RNA polymerase (dark dots along the DNA), moving left to right (red arrow). The growing RNA transcripts extend from the DNA, accounting for the feather-like appearance. In the diagram of a transcription unit below the TEM, the genes for three types of rRNA (darker blue) are adjacent to regions that are transcribed but later removed (medium blue). A single transcript is processed to yield one of each of the three rRNAs (red), key components of the ribosome.



(b) The human α-globin and β-globin gene families. Adult hemoglobin is composed of two α-globin and two β-globin polypeptide subunits, as shown in the molecular model. The genes (darker blue) encoding α- and β-globins are found in two families, organized as shown here. The noncoding DNA (light blue) separating the functional genes within each family includes pseudogenes (ψ; gold), versions of the functional genes that no longer encode functional polypeptides. Genes and pseudogenes are named with Greek letters, as you have seen previously for the α- and β-globins. Some genes are expressed only in the embryo or fetus.

VISUAL SKILLS ➤ *In the TEM at the top of part (a), how could you determine the direction of transcription if it weren't indicated by the red arrow?*

CONCEPT 20.5

Duplication, rearrangement, and mutation of DNA contribute to genome evolution

EVOLUTION Now that we have explored the makeup of the human genome, let's see what its composition reveals about how the genome evolved. The basis of change at the genomic level is mutation, which underlies much of genome evolution. It seems likely that the earliest forms of life had a minimal number of genes—those necessary for survival and reproduction. If this were indeed the case, one aspect of evolution must have been an increase in the size of the genome, with the extra genetic material providing the raw material for gene diversification. In this section, we'll first describe how extra copies of all or part of a genome can arise and then consider subsequent processes that can lead to the evolution of proteins (or RNA products) with slightly different or entirely new functions.

Duplication of Entire Chromosome Sets

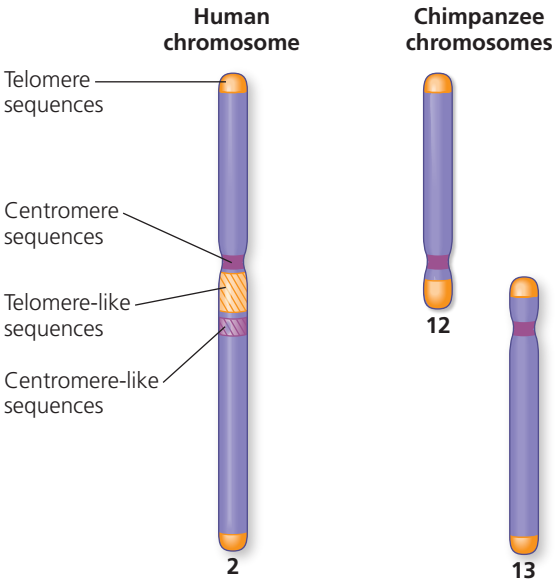
An accident in meiosis, such as failure to separate homologs during meiosis I, can result in one or more extra sets of chromosomes, a condition known as polyploidy (see Concept 15.4). Although such accidents would most often be lethal, in rare cases they could facilitate the evolution of genes. In a polyploid organism, one set of genes can provide essential functions for the organism. The genes in the one or more extra sets can diverge by accumulating mutations; these variations may persist if the organism carrying them survives and reproduces. In this way, genes with novel functions can evolve. As long as one copy of an essential gene is expressed, the divergence of another copy can lead to its encoded protein acting in a novel way, thereby changing the organism's phenotype.

The outcome of this accumulation of mutations may eventually be the branching off of a new species. While polyploidy is rare among animals, it is relatively common among plants, especially flowering plants. Some botanists estimate that as many as 80% of the plant species that are alive today show evidence of polyploidy having occurred among their ancestral species. You'll learn more about how polyploidy leads to plant speciation in Concept 24.2.

Alterations of Chromosome Structure

With the recent explosion in genomic sequence information, we can now compare the chromosomal organizations of many different species in detail. This information allows us to make inferences about the evolutionary processes that shape chromosomes and may drive speciation. For example, scientists have long known that sometime in the last 6 million years, when the ancestors of humans and chimpanzees diverged as species, the fusion of two ancestral chromosomes in the human line led to different haploid numbers for humans ($n = 23$) and chimpanzees ($n = 24$). The banding patterns in stained chromosomes suggested that the ancestral versions of current chimpanzee chromosomes 12 and 13 fused end to end, forming chromosome 2 in an ancestor of the human lineage. Sequencing and analysis of human chromosome 2 during the Human Genome Project provided very strong supporting evidence for the model we have just described **(Figure 20.11)**.

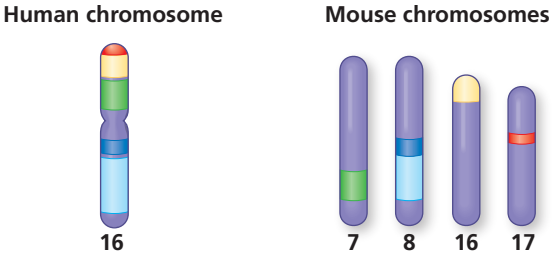
Figure 20.11 Human and chimpanzee chromosomes. The positions of telomere-like and centromere-like sequences on human chromosome 2 (left) match those of telomeres on chimpanzee chromosomes 12 and 13 and the centromere on chimpanzee chromosome 13 (right). This suggests that chromosomes 12 and 13 in a human ancestor fused end to end to form human chromosome 2. The centromere from ancestral chromosome 12 remained functional on human chromosome 2, while the one from ancestral chromosome 13 did not.



In another study of broader scope, researchers compared the DNA sequence of each human chromosome with the whole-genome sequence of the mouse **(Figure 20.12)**. One part of their study showed that large blocks of genes on human chromosome 16 are found on four mouse chromosomes, indicating that the genes in each block stayed together in both the mouse and the human lineages during their divergent evolution from a common ancestor.

Performing the same comparison of chromosomes of humans and six other mammalian species allowed the researchers to reconstruct the evolutionary history of chromosomal rearrangements in these eight species. They found many duplications and inversions of large portions of chromosomes, the result of errors during meiotic recombination in which the DNA was broken and rejoined incorrectly. The rate of these events seems to have begun accelerating about 100 million years ago, around 35 million years before large

Figure 20.12 Human and mouse chromosomes. Here, we can see that DNA sequences very similar to large blocks of human chromosome 16 (colored areas in this diagram) are found on mouse chromosomes 7, 8, 16, and 17. This finding suggests that the DNA sequence in each block has stayed together in the mouse and human lineages since the time they diverged from a common ancestor.



dinosaurs became extinct and the number of mammalian species began rapidly increasing. The apparent coincidence is interesting because chromosomal rearrangements are thought to contribute to the generation of new species. Although two individuals with different arrangements could still mate and produce offspring, the offspring would have two nonequivalent sets of chromosomes, making meiosis inefficient or even impossible. Thus, chromosomal rearrangements would lead to two populations that could not successfully mate with each other, a step on the way to their becoming two separate species. (You'll learn more about this in Concept 24.2.)

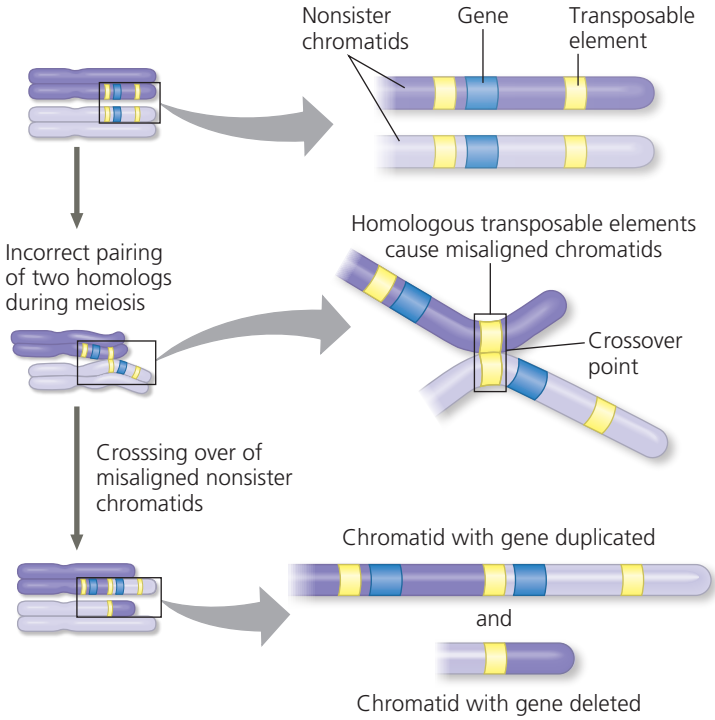
The same study also unearthed a pattern with medical relevance. Analysis of the chromosomal breakage points associated with the rearrangements showed that specific sites were used over and over again. A number of these recombination "hot spots" correspond to locations of chromosomal rearrangements within the human genome that are associated with congenital diseases (see Concept 15.4).

Duplication and Divergence of Gene-Sized Regions of DNA

Errors during meiosis can also lead to the duplication of chromosomal regions that are smaller than the ones we've just discussed, including segments the length of individual genes. Unequal crossing over during prophase I of meiosis, for instance, can result in one chromosome with a deletion and another with a duplication of a particular gene. Transposable elements can provide homologous sites where nonsister chromatids can cross over, even when other chromatid regions are not correctly aligned **(Figure 20.13)**.

Also, slippage can occur during DNA replication, such that the template shifts with respect to the new complementary strand, and a part of the template strand is either skipped by the replication machinery or used twice as a template. As a result, a segment of DNA is deleted or duplicated. It is easy to imagine how such errors could occur in regions of repeats. The variable number of repeated units of simple sequence

Figure 20.13 Gene duplication due to unequal crossing over. One mechanism by which a gene (or other DNA segment) can be duplicated is recombination during meiosis between copies of a transposable element (yellow) flanking the gene (blue). Such recombination between misaligned nonsister chromatids of homologous chromosomes produces one chromatid with two copies of the gene and one chromatid with no copy. (Genes and transposable elements are shown only in the region of interest.)



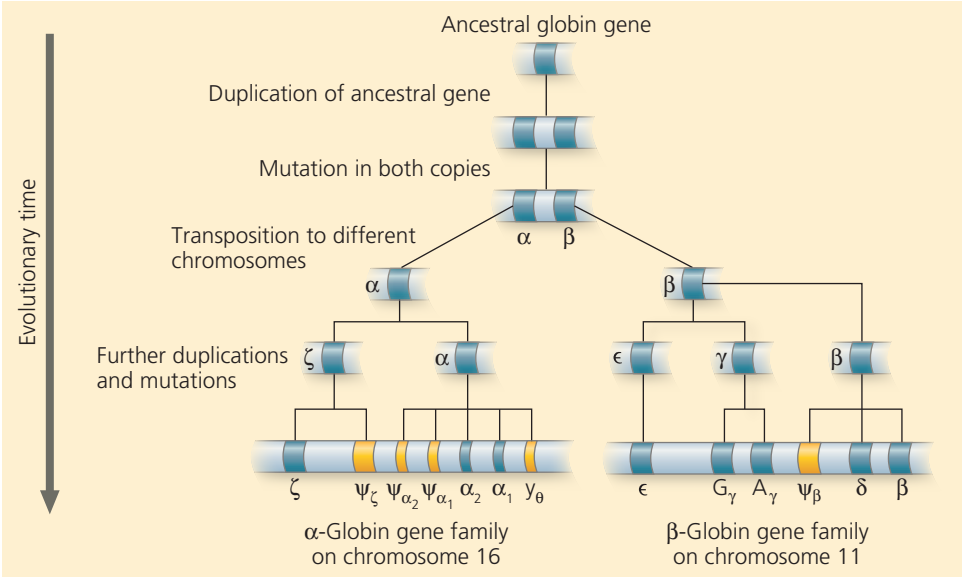
MAKE CONNECTIONS ➤ *Examine how crossing over occurs in Figure 13.9. In the middle panel above, draw a line through the portions that result in the upper chromatid in the bottom panel. Use a different color to do the same for the other chromatid.*

DNA at a given site, used for STR analysis, is probably due to errors like these. Evidence that unequal crossing over and template slippage during DNA replication lead to duplication of genes is found in the existence of multigene families, such as the globin family.

Evolution of Genes with Related Functions: The Human Globin Genes

In Figure 20.10b, you saw the organization of the α-globin and β-globin gene families as they exist in the human genome today. Now, let's consider how events such as duplications can lead to the evolution of genes with related functions like the globin genes. A comparison of gene sequences within a multigene family can suggest the order in which the genes arose. Re-creating the evolutionary history of the globin genes using this approach indicates that they all evolved from one common ancestral globin gene that underwent duplication and divergence into the α-globin and β-globin ancestral genes about 450–500 million years ago. Each of these genes was later duplicated several times, and the copies then diverged from each other in sequence, yielding the

Figure 20.14 A proposed model for the sequence of events in the evolution of the human α-globin and β-globin gene families from a single ancestral globin gene.



? *The gold elements are pseudogenes. Explain how they could have arisen after gene duplication.*

current family members **(Figure 20.14)**. In fact, the common ancestral globin gene also gave rise to the oxygen-binding muscle protein myoglobin and to the plant protein leghemoglobin. The latter two proteins function as monomers, and their genes are included in a "globin superfamily."

After the duplication events, the differences between the genes in the globin families undoubtedly arose from mutations that accumulated in the gene copies over many generations. The current model is that the necessary function provided by an α-globin protein, for example, was fulfilled by one gene, while other copies of the α-globin gene accumulated random mutations. Many mutations may have had an adverse effect on the organism, and others may have had no effect. However, a few mutations must have altered the function of the protein product in a way that benefitted the organism at a particular life stage without substantially changing the protein's oxygen-carrying function. Presumably, natural selection acted on these altered genes, maintaining them in the population.

In the **Scientific Skills Exercise**, you can compare amino acid sequences of the globin family proteins and see how such comparisons were used to generate the model for globin gene evolution shown in Figure 20.14. The existence of several pseudogenes among the functional globin genes provides additional evidence for this model: Random mutations in these "genes" over evolutionary time have destroyed their function.

Evolution of Genes with Novel Functions

In the evolution of the globin gene families, gene duplication and subsequent divergence produced family members whose protein products performed functions similar to each other (oxygen transport). However, an alternative scenario is that one copy of a duplicated gene can undergo alterations that lead to a completely new function for the protein product. The genes for lysozyme and α-lactalbumin are a good example.

Lysozyme is an enzyme that helps protect animals against bacterial infection by hydrolyzing bacterial cell walls (see Visualizing Figure 5.16); α-lactalbumin is a nonenzymatic protein that plays a role in milk production in mammals. The two proteins are quite similar in their amino acid sequences and three-dimensional structures **(Figure 20.15)**. Both genes are found in mammals, but only the lysozyme gene is present in birds. These findings suggest that at some time after the lineages leading to mammals and birds had separated, the lysozyme gene was duplicated in the mammalian lineage but not in the avian lineage. Subsequently, one copy of the duplicated lysozyme gene evolved into a gene encoding α-lactalbumin, a protein with a completely new function associated with a key characteristic of mammals—milk production. In a recent study, evolutionary biologists searched vertebrate genomes for genes with similar sequences. There appear to be at least eight members of the lysozyme family, distributed widely among mammalian species. The functions of all the encoded gene products are not yet known, but it will be exciting to discover whether they are as different as the functions of lysozyme and α-lactalbumin.

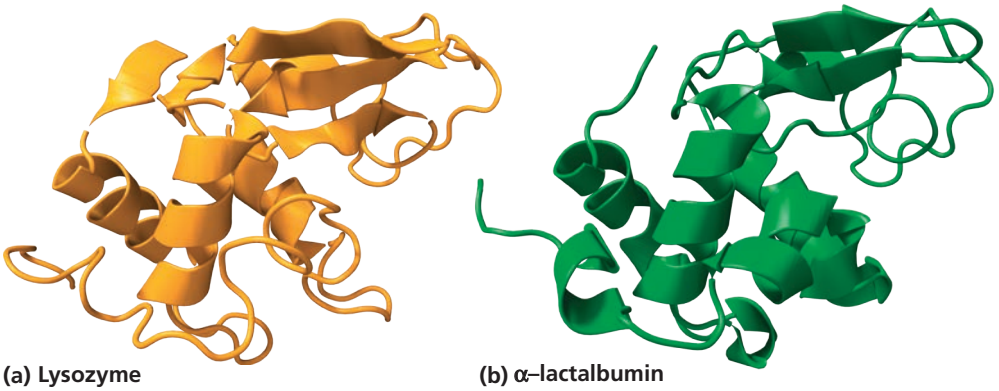
Besides the duplication and divergence of whole genes, rearrangement of existing DNA sequences within genes has also contributed to genome evolution. The presence of introns may have promoted the evolution of new proteins by facilitating the duplication or shuffling of exons, as we'll discuss next.

Rearrangements of Parts of Genes: Exon Duplication and Exon Shuffling

Recall from Concept 17.3 that an exon often codes for a protein domain, a distinct structural and functional region of a protein molecule. We've already seen that unequal crossing over during meiosis can lead to duplication of a gene on one chromosome and its loss from the homologous chromosome (see Figure 20.13). By a similar process, a particular exon within a gene could be duplicated on one chromosome and deleted from the other. The gene with the duplicated exon would code for a protein containing a second copy

Figure 20.15 Comparison of lysozyme and α-lactalbumin proteins.
Computer-generated ribbon models of the similar structures of **(a)** lysozyme and **(b)** α-lactalbumin are shown, along with **(c)** a comparison of the amino acid sequences of the two proteins. The amino acids are arranged in groups of ten for ease of reading, and single-letter amino acid codes are used (see Figure 5.14). Identical amino acids are highlighted in yellow, and dashes indicate gaps in one sequence that have been introduced by the software to optimize the alignment.

MAKE CONNECTIONS ➤ *Even though two amino acids are not identical, they may be structurally and chemically similar and therefore behave similarly. Using Figure 5.14 as a reference, examine the nonidentical amino acids in positions 1–30 and note cases where the amino acids in the two sequences are similarly acidic or basic.*



(a) Lysozyme (b) α–lactalbumin

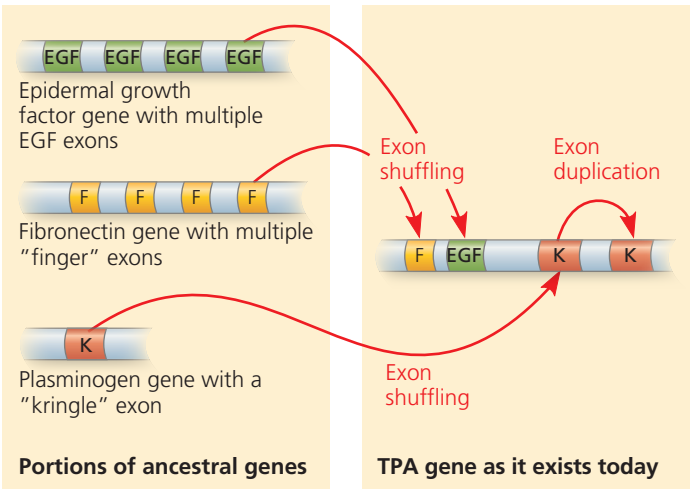
Lysozyme	**1**	KVFERCELAR	TLKRLGMDGY	RGISLANWMC	LAKWESGYNT	RATNYNAGDR
α–lactalbumin	**1**	KQFTKCELSQ	LLK--DIDGY	GGIALPELIC	TMFHTSGYDT	QAIVENN--E
Lysozyme	**51**	STDYGIFQIN	SRYWCNDGKT	PGAVNACHLS	CSALLQDNIA	DAVACAKRVV
α–lactalbumin	**51**	STEYGLFQIS	NKLWCKSSQV	PQSRNICDIS	CDKFLDDDIT	DDIMCAKKIL
Lysozyme	**101**	RDPQGIRAWV	AWRNRCQ-NR	DVRQYVQGCG	V
α–lactalbumin	**101**	D-IKGIDYWL	AHKALCT--E	KLEQWLCEKL	–

(c) Amino acid sequence alignments of lysozyme and α–lactalbumin

of the encoded domain. This change in the protein's structure might augment its function by increasing its stability, enhancing its ability to bind a particular ligand, or altering some other property. Quite a few protein-coding genes have multiple copies of related exons, which presumably arose by duplication and then diverged. The gene encoding the extracellular matrix protein collagen is a good example. Collagen is a structural protein (see Figure 5.18) with a highly repetitive amino acid sequence, which reflects the repetitive pattern of exons in the collagen gene.

As an alternative possibility, we can imagine the occasional mixing and matching of different exons either within a gene or between two different (nonallelic) genes owing to errors in meiotic recombination. This process, termed *exon shuffling*, could lead to new proteins with novel combinations of functions. As an example, let's consider the gene for tissue plasminogen activator (TPA). The TPA protein is an extracellular protein that helps control blood clotting. It has four domains of three types, each encoded by an exon, and one of those exons is present in two copies. Because each type of exon is also found in other proteins, the current version of the gene for TPA is thought to have arisen by several instances of exon shuffling during errors in meiotic recombination and subsequent duplication **(Figure 20.16)**.

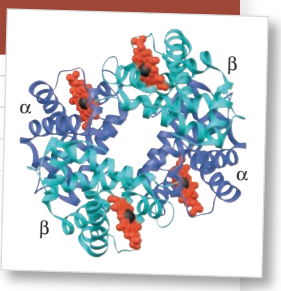
Figure 20.16 Evolution of a new gene by exon shuffling.
Meiotic errors could have moved exons, each encoding a particular domain, from ancestral forms of the genes for epidermal growth factor, fibronectin, and plasminogen (left) into the evolving gene for tissue plasminogen activator, TPA (right). Subsequent duplication of the "kringle" exon (K) from the plasminogen gene after its movement into the TPA gene could account for the two copies of this exon in the TPA gene existing today.



VISUAL SKILLS ➤ *Looking at Figure 20.13, describe the steps by which transposable elements within introns might have facilitated the exon shuffling shown here.*

SCIENTIFIC SKILLS EXERCISE

Reading an Amino Acid Sequence Identity Table



Hemoglobin

How Have Amino Acid Sequences of Human Globin Genes Diverged During Their Evolution? To build a model of the evolutionary history of the globin genes (see Figure 20.14), researchers compared the amino acid sequences of the polypeptides they encode. In this exercise, you will analyze comparisons of the amino acid sequences of the globin polypeptides to shed light on their evolutionary relationships.

How the Experiment Was Done Scientists obtained the DNA sequences for each of the eight globin genes and "translated" them into amino acid sequences. They then used a computer program to align the sequences (with dashes indicating gaps in one sequence) and calculate a percent identity value for each pair of globins. The percent identity reflects the number of positions with identical amino acids relative to the total number of amino acids in a globin polypeptide. The data were displayed in a table to show the pairwise comparisons.

Data from the Experiment The following table shows an example of a pairwise alignment—that of the α_1-globin (alpha-1 globin) and ζ-globin (zeta globin) amino acid sequences—using the standard single-letter symbols for amino acids. To the left of each line of amino acid sequence is the number of the first amino acid in that line. The percent identity value for the α_1- and ζ-globin amino acid sequences was calculated by counting the number of matching amino acids (86, highlighted in yellow), dividing by the total number of amino acid positions (143), and then multiplying by 100. This resulted in a 60% identity value for the α_1-ζ pair, as shown in the amino acid identity table at the bottom of the page. The values for other globin pairs were calculated in the same way.

Globin	Alignment of Globin Amino Acid Sequences
α_1	1 MVLSPADKTNVKAAWGKVGAHAGEYGAEAL
ζ	1 MSLTKTERTIIVSMWAKISTQADTIGTETL
α_1	31 ERMFLSFPTTKTYFPHFDLSH-GSAQVKGH
ζ	31 ERLFLSHPQTKTYFPHFDL-HPGSAQLRAH
α_1	61 GKKVADALTNAVAHVDDMPNALSALSDLHA
ζ	61 GSKVVAAVGDAVKSIDDIGGALSKLSELHA
α_1	91 HKLRVDPVNFKLLSHCLLVTLAAHLPAEFT
ζ	91 YILRVDPVNFKLLSHCLLVTLAARFPADFT
α_1	121 PAVHASLDKFLASVSTVLTSKYR
ζ	121 AEAHAAWDKFLSVVSSVLTEKYR

INTERPRET THE DATA

1. Note that in the amino acid identity table, the data are arranged so each globin pair can be compared. (a) Some cells in the table have dashed lines. Given the pairs that are being compared for these cells, what percent identity value is implied by the dashed lines? (b) Notice that the cells in the lower left half of the table are blank. Using the information already provided in the table, fill in the missing values. Why does it make sense that these cells were left blank in the table?
2. The earlier that two genes arose from a duplicated gene, the more their nucleotide sequences can have diverged, which may result in amino acid differences in the protein products. (a) Based on that premise, identify which two genes are most divergent from each other. What is the percent amino acid identity between their polypeptides? (b) Using the same approach, identify which two globin genes are the most recently duplicated. What is the percent identity between them?
3. The model of globin gene evolution shown in Figure 20.14 suggests that an ancestral gene duplicated and mutated to become α- and β-globin genes, and then each one was further duplicated and mutated. What features of the data set support the model?
4. Make an ordered list of all the percent identity values from the table, starting with 100% at the top. Next to each number write the globin pair(s) with that percent identity value. Use one color for the globins from the α family and a different color for the globins from the β family. (a) Compare the order of pairs on your list with their positions in the model shown in Figure 20.14. Does the order of pairs describe the same relative "closeness" of globin family members seen in the model? (b) Compare the percent identity values for pairs within the α or β group to the values for between-group pairs.

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Further Reading R. C. Hardison, Globin genes on the move, *Journal of Biology* 7:35.1–35.5 (2008).

Amino Acid Identity Table

		α Family			β Family				
		α_1 (alpha 1)	α_2 (alpha 2)	ζ (zeta)	β (beta)	δ (delta)	ϵ (epsilon)	A_γ (gamma A)	G_γ (gamma G)
α Family	α_1	-----	100	60	45	44	39	42	42
	α_2		-----	60	45	44	39	42	42
	ζ			-----	38	40	41	41	41
β Family	β				-----	93	76	73	73
	δ					-----	73	71	72
	ϵ						-----	80	80
	A_γ							-----	99
	G_γ								-----

Compiled using data from the National Center for Biotechnology Information (NCBI).

How Transposable Elements Contribute to Genome Evolution

The persistence of transposable elements as a large fraction of some eukaryotic genomes is consistent with the idea that they play an important role in shaping a genome over evolutionary time. These elements can contribute to the evolution of the genome in several ways. They can promote recombination, disrupt cellular genes or control elements, and carry entire genes or individual exons to new locations.

Transposable elements of similar sequence scattered throughout the genome facilitate recombination between different (nonhomologous) chromosomes by providing homologous regions for crossing over (see Figure 20.13). Most such recombination events are probably detrimental, causing chromosomal translocations and other changes in the genome that may be lethal to the organism. But over the course of evolutionary time, an occasional recombination event of this sort may be advantageous to the organism. (For the change to be heritable, of course, it must happen in a cell that will give rise to a gamete.)

The movement of a transposable element can have a variety of consequences. For instance, a transposable element that "jumps" into a protein-coding sequence will prevent the production of a normal transcript of the gene. (Introns provide a sort of "safety zone" that does not affect the transcript because the transposable element will be spliced out.) If a transposable element inserts within a regulatory sequence, the transposition may lead to increased or decreased production of one or more proteins. Transposition caused both types of effects on the genes coding for pigment-synthesizing enzymes in McClintock's corn kernels. Again, while such changes are usually harmful, in the long run some may provide a survival advantage. A possible example was mentioned earlier: At least some of the *Alu* transposable elements in the human genome are known to produce RNAs that regulate expression of human genes.

During transposition, a transposable element may carry along a gene or even a group of genes to a new position in the genome. This mechanism probably accounts for the location of the α-globin and β-globin gene families on different human chromosomes, as well as the dispersion of the genes of certain other gene families. By a similar tag-along process, an exon from one gene may be inserted into another gene in a mechanism similar to that of exon shuffling during recombination. For example, an exon may be inserted by transposition into the intron of a protein-coding gene. If the inserted exon is retained in the RNA transcript during RNA splicing, the protein that is synthesized will have an additional domain, which may confer a new function on the protein.

Most often, the processes discussed in this section produce harmful effects, which may be lethal, or have no effect at all. In a few cases, however, small heritable changes that are beneficial may occur. Over many generations, the resulting genetic diversity provides valuable raw material for natural selection. Diversification of genes and their products is an important factor in the evolution of new species. Thus, the accumulation of changes in the genome of each species provides a record of its evolutionary history. To read this record, we must be able to identify genomic changes. Comparing the genomes of different species allows us to do that, increasing our understanding of how genomes evolve. You will learn more about these topics next.

CONCEPT CHECK 20.5

1. Describe three examples of errors in cellular processes that lead to DNA duplications.
2. Explain how multiple exons might have arisen in the ancestral EGF and fibronectin genes shown on the left side of Figure 20.16.
3. What are three ways that transposable elements are thought to contribute to genome evolution?
4. **WHAT IF? ➤** In 2005, Icelandic scientists reported finding a large chromosomal inversion present in 20% of northern Europeans, and they noted that Icelandic women with this inversion had significantly more children than women without it. What would you expect to happen to the frequency of this inversion in the Icelandic population in future generations?

For suggested answers, see Appendix A.

CONCEPT 20.6

Comparing genome sequences provides clues to evolution and development

EVOLUTION One researcher has likened the current state of biology to the Age of Exploration in the 1400s, which occurred soon after major improvements in navigation and ship design. In the last 30 years, we have seen rapid advances in genome sequencing and data collection, new techniques for assessing gene activity across the whole genome, and refined approaches for understanding how genes and their products work together in complex systems. In the field of biology, we are truly poised on the brink of a new world.

Comparisons of genome sequences from different species reveal a lot about the evolutionary history of life, from very ancient to more recent. Similarly, comparative studies of the genetic programs that direct embryonic development in different species are beginning to clarify the mechanisms that generated the great diversity of life-forms present today. In this final section of the chapter, we will discuss what has been learned from these two approaches.

Comparing Genomes

The more similar in sequence the genes and genomes of two species are, the more closely related those species are in their

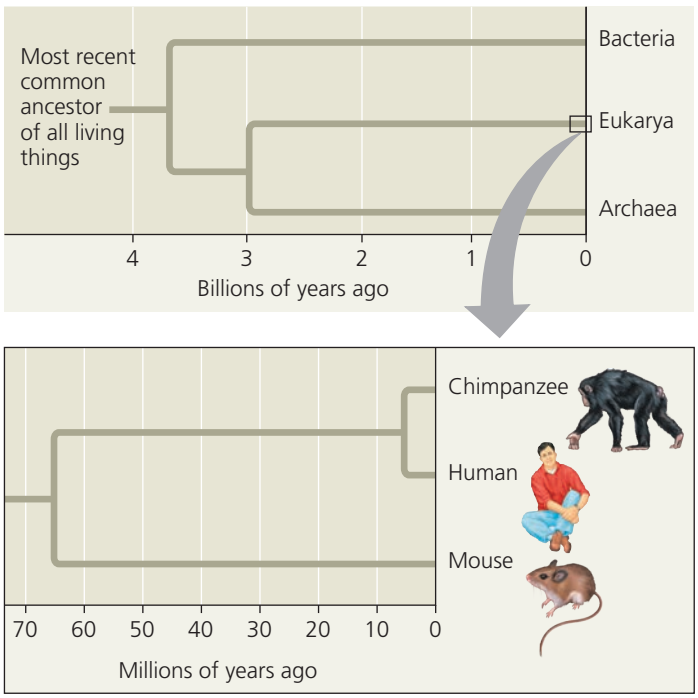
evolutionary history—because not enough time has passed for mutations and other changes to accumulate. Comparing genomes of closely related species sheds light on more recent evolutionary events, whereas comparing genomes of very distantly related species helps us understand ancient evolutionary history. In either case, learning about characteristics that are shared or divergent between groups enhances our picture of the evolution of organisms and biological processes. As you learned in Concept 1.2, the evolutionary relationships between species can be represented by a diagram in the form of a tree, where each branch point marks the divergence of two lineages. **Figure 20.17** shows the evolutionary relationships of some groups and species we will discuss.

Comparing Distantly Related Species

Determining which genes have remained similar—that is, are *highly conserved*—in distantly related species can help clarify evolutionary relationships among species that diverged from each other long ago. Indeed, comparisons of the specific gene sequences of bacteria, archaea, and eukaryotes indicate that these three groups diverged between 2 and 4 billion years ago and strongly support the theory that they are the fundamental domains of life (see Figure 20.17).

In addition to their value in evolutionary biology, comparative genomic studies confirm the relevance of research on model organisms to our understanding of biology in general and human biology in particular. Very ancient genes can still be surprisingly similar in disparate species. A 2015 study tested the ability of the human version of each of 414 important yeast genes to function equivalently in yeast cells. Remarkably, the researchers concluded that 47% of these yeast genes could be replaced by the human gene. This striking result underscores the common origin of yeasts and humans—two distantly related species.

Figure 20.17 Evolutionary relationships of the three domains of life. The tree diagram at the top shows the ancient divergence of bacteria, archaea, and eukaryotes. A portion of the eukaryote lineage is expanded in the inset to show the more recent divergence of three mammalian species discussed in this chapter.



Comparing Closely Related Species

The genomes of two closely related species are likely to be organized similarly because of their relatively recent divergence. Their long shared history also means that only a small number of gene differences are found when their genomes are compared. These genetic differences can thus be more easily correlated with phenotypic differences between the two species. An exciting application of this type of analysis is seen as researchers compare the human genome with the genomes of the chimpanzee, mouse, rat, and other mammals. Identifying the genes shared by all of these species but not by nonmammals gives us clues about what it takes to make a mammal, while finding the genes shared by chimpanzees and humans but not by rodents tells us something about primates. And, of course, comparing the human genome with that of the chimpanzee helps us answer the tantalizing question we asked at the beginning of the chapter: What genomic information defines a human or a chimpanzee?

An analysis of the overall composition of the human and chimpanzee genomes, which are thought to have diverged only about 6 million years ago (see Figure 20.17), reveals some general differences. Considering single nucleotide substitutions, the two genomes differ by only 1.2%. When researchers looked at longer stretches of DNA, however, they were surprised to find a further 2.7% difference due to insertions or deletions of larger regions in the genome of one or the other species; many of the insertions were duplications or other repetitive DNA. In fact, a third of the human duplications are not present in the chimpanzee genome, and some of these duplications contain regions associated with human diseases. There are more *Alu* elements in the human genome than in the chimpanzee genome, and the latter contains many copies of a retroviral provirus (see Chapter 26) not present in humans. All of these observations provide clues to the forces that might have swept the two genomes along different paths, but we don't have a complete picture yet.

Along with chimpanzees, bonobos are the other African ape species that are the closest living relatives to humans. The sequencing of the bonobo genome, completed in 2012, revealed that in some regions, human sequences were more closely related to either chimpanzee or bonobo sequences than chimpanzee or bonobo sequences were to each other. Such a fine-grained comparison of three closely related species allows even more detail to be worked out in reconstructing their related evolutionary history.

We also don't know how the genetic differences revealed by genome sequencing might account for the distinct characteristics of each species. To discover the basis for the phenotypic differences between chimpanzees and humans, biologists are studying specific genes and types of genes that differ between the two species and comparing them with their counterparts in other mammals. This approach has revealed a number of genes that are apparently changing (evolving) faster in the human than in either the chimpanzee or the mouse. Among them are genes involved in defense against malaria and tuberculosis as well as at least one gene that regulates brain size. When genes are classified by function, the genes that seem to be evolving the fastest are those that code for transcription factors. This discovery makes sense because transcription factors regulate gene expression and thus play a key role in orchestrating the overall genetic program.

One transcription factor whose gene shows evidence of rapid change in the human lineage is called FOXP2 **(Figure 20.18)**. Several lines of evidence suggest that the *FOXP2* gene product regulates genes that function in

Figure 20.18

Inquiry What is the function of a gene (*FOXP2*) that is rapidly evolving in the human lineage?

Experiment Several lines of evidence support a role for the *FOXP2* gene in the development of speech and language in humans and of vocalization in other vertebrates. In 2005, Joseph Buxbaum and collaborators at the Mount Sinai School of Medicine and several other institutions tested the function of *FOXP2*. They used the mouse, a model organism in which genes can be easily knocked out, as a representative vertebrate that vocalizes: Mice produce ultrasonic squeaks (whistles) to communicate stress. The researchers used genetic engineering to produce mice in which one or both copies of *FOXP2* were disrupted.

Wild type: two normal copies of *FOXP2*

Heterozygote: one copy of *FOXP2* disrupted

Homozygote: both copies of *FOXP2* disrupted

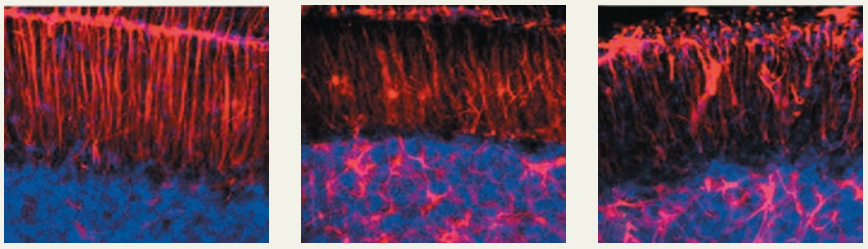
They then compared the phenotypes of these mice. Two of the characters they examined are included here: brain anatomy and vocalization.

Experiment 1: Researchers cut thin sections of brain and stained them with reagents that allow visualization of brain anatomy in a UV fluorescence microscope.

Experiment 2: To induce stress, researchers separated each newborn mouse pup from its mother and recorded the number of ultrasonic whistles produced by the pup.

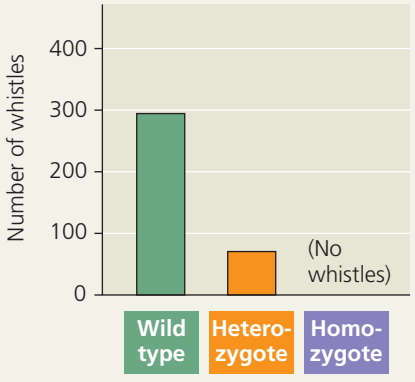
Results

Experiment 1 Results: Disruption of both copies of *FOXP2* led to brain abnormalities in which the cells were disorganized. Phenotypic effects on the brain of heterozygotes, with one disrupted copy, were less severe. (Each color in the micrographs below reveals a different cell or tissue type.)



Wild type | Heterozygote | Homozygote

Experiment 2 Results: Disruption of both copies of *FOXP2* led to an absence of ultrasonic vocalization in response to stress. The effect on vocalization in the heterozygote was also extreme.



Conclusion *FOXP2* plays a significant role in the development of functional communication systems in mice. The results augment evidence from studies of birds and humans, supporting the hypothesis that *FOXP2* may act similarly in diverse organisms.

Data from W. Shu et al., Altered ultrasonic vocalization in mice with a disruption in the *Foxp2* gene, *Proceedings of the National Academy of Sciences USA* 102:9643–9648 (2005).

WHAT IF? ➤ *Since the results support a role for mouse FOXP2 in vocalization, you might wonder whether the human FOXP2 protein is a key regulator of speech. If you were given the amino acid sequences of wild-type and mutant human FOXP2 proteins and the wild-type chimpanzee FOXP2 protein, how would you investigate this question? What further clues could you obtain by comparing these sequences to that of the mouse FOXP2 protein?*

vocalization in vertebrates. First, mutations in this gene can produce severe speech and language impairment in humans. Moreover, the *FOXP2* gene is expressed in the brains of zebra finches and canaries at the time when these songbirds are learning their songs. But perhaps the strongest evidence comes from a "knockout" experiment in which researchers disrupted the *FOXP2* gene in mice and analyzed the resulting phenotype (see Figure 20.18). The homozygous mutant mice had malformed brains and failed to emit normal ultrasonic vocalizations, and mice with one faulty copy of the gene also showed significant problems with vocalization. These results support the idea that the *FOXP2* gene product turns on genes involved in vocalization.

Expanding on this analysis, another research group more recently replaced the *FOXP2* gene in mice with a "humanized" copy coding for the human versions of two amino acids that differ between human and chimpanzee; these are the changes potentially responsible for a human's ability to speak. Although the mice were generally healthy, they had subtly different vocalizations and showed changes in brain cells in circuits associated with speech in human brains.

In 2010, the Neanderthal genome was sequenced from a very small amount of preserved genomic DNA, and a high-quality sequence was completed in 2014. Neanderthals (*Homo neanderthalensis*) are members of the same genus to which humans (*Homo sapiens*) belong (see Concept 34.7). A reconstruction of their evolutionary history based on genomic comparisons between the two species suggests that some groups of humans and Neanderthals co-existed and interbred for a period of time before Neanderthals went extinct about 30,000 years ago. While Neanderthals have sometimes been portrayed as primitive beings that could only grunt, their *FOXP2* gene sequence encodes a protein identical to that of humans. This suggests that Neanderthals may have been capable of speech of some type and, along with other observed genetic similarities, forces us to reevaluate our image of our recent extinct relatives.

The *FOXP2* story is an excellent example of how different approaches can complement each other in uncovering biological phenomena of widespread importance. The *FOXP2* experiments used mice as a model for humans because it would be unethical (as well as impractical) to carry out such experiments in humans. Mice and humans, which diverged about 65.5 million years ago (see Figure 20.17), share about 85% of their genes. This genetic similarity can be exploited in studying human genetic disorders. If researchers know the organ or tissue that is affected by a particular genetic disorder, they can look for genes that are expressed in these locations in mice.

Even though more distantly related to humans, fruit flies have also been a useful model species for study of such human disorders as Parkinson's disease and alcoholism, while nematodes (soil worms) have yielded a wealth of information about aging. Further research efforts are under way to extend genomic studies to many more species, including neglected species from diverse branches of the tree of life. These studies will advance our understanding of evolution, of course, as well as all aspects of biology, from human health to ecology.

Comparing Genomes Within a Species

Another exciting consequence of our ability to analyze genomes is our growing understanding of the spectrum of genetic variation in humans. Because the history of the human species is so short—probably about 200,000 years—the amount of DNA variation among humans is small compared to that of many other species. Much of our diversity seems to be in the form of single nucleotide polymorphisms (SNPs). SNPs are single base-pair sites where genetic variation is found in at least 1% of the population (see Concept 19.2); they are usually detected by DNA sequencing. In the human genome, SNPs occur on average about once in 100–300 base pairs. Scientists have already identified the locations of several million SNP sites in the human genome and continue to find additional SNPs. These are stored in databases around the world, one of which is run by the National Center for Biotechnology Information (NCBI) and can be accessed at http://www.ncbi.nlm.nih.gov/SNP/.

In the course of this search, they have also found other variations—including chromosomal regions with inversions, deletions, and duplications. The most surprising discovery has been the widespread occurrence of *copy-number variants* (*CNVs*), loci where some individuals have one or multiple copies of a particular gene or genetic region rather than the standard two copies (one on each homolog). CNVs result from regions of the genome being duplicated or deleted inconsistently within the population. One study of 40 people found more than 8,000 CNVs involving 13% of the genes in the genome, and these CNVs probably represent just a small subset of the total. Since these variants encompass much longer stretches of DNA than the single nucleotides of SNPs, CNVs are more likely to have phenotypic consequences and to play a role in complex diseases and disorders. At the very least, the high incidence of copy-number variation blurs the meaning of the phrase "a normal human genome."

Copy-number variants, SNPs, and variations in repetitive DNA such as short tandem repeats (STRs) are useful genetic markers for studying human evolution. In one study, the genomes of two Africans from different communities were sequenced: Archbishop Desmond Tutu, the South African civil rights advocate and a member of the Bantu tribe, the majority population in southern Africa, and !Gubi, a hunter-gatherer from the Khoisan community in Namibia, a minority African population that is probably the human group with the oldest known lineage. The comparison revealed many differences, as you might expect. The analysis was then broadened to compare the protein-coding regions of !Gubi's genome with those of three other

Khoisan community members (self-identified Bushmen) living nearby. Remarkably, the four African genomes differed more from each other than a European would from an Asian. These data highlight the extensive diversity among African genomes. Extending this approach will help us answer important questions about the differences between human populations and the migratory routes of human populations throughout history.

Widespread Conservation of Developmental Genes Among Animals

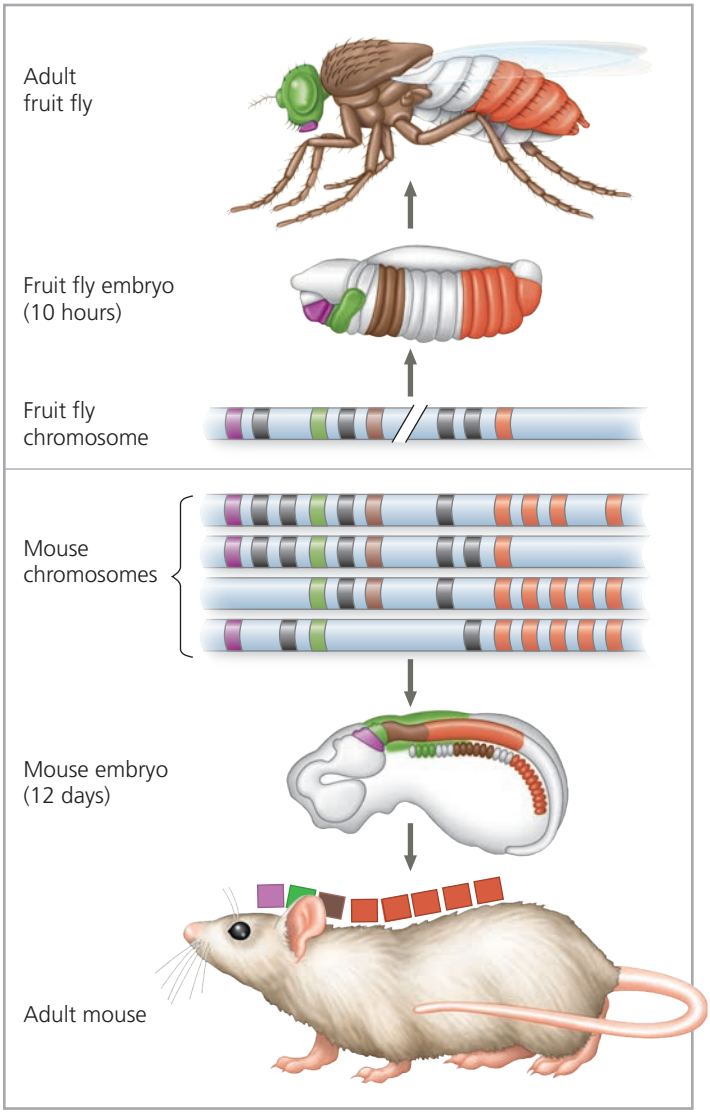
Biologists in the field of evolutionary developmental biology, or **evo-devo** as it is often called, compare developmental processes of different multicellular organisms. Their aim is to understand how these processes have evolved and how changes in them can modify existing organismal features or lead to new ones. With the advent of molecular techniques and the recent flood of genomic information, we are beginning to realize that the genomes of related species with strikingly different forms may have only minor differences in gene sequence or, perhaps more importantly, in gene regulation. Discovering the molecular basis of these differences in turn helps us understand the origins of the myriad diverse forms that cohabit this planet, thus informing our study of the evolution of life.

In Concept 18.4, you learned about the homeotic genes in *Drosophila melanogaster* (see Figure 18.20), which encode transcription factors that regulate gene expression and specify the identity of body segments in the fruit fly. Molecular analysis of the homeotic genes in *Drosophila* has shown that they all include a 180-nucleotide sequence called a **homeobox**, which codes for a 60-amino-acid *homeodomain* in the encoded proteins. An identical or very similar nucleotide sequence has been discovered in the homeotic genes of many invertebrates and vertebrates. In fact, the nucleotide sequences in humans and fruit flies are so similar that one researcher has whimsically referred to flies as "little people with wings." The resemblance even extends to the organization of these genes: The vertebrate genes homologous to the homeotic genes of fruit flies have kept the same chromosomal arrangement **(Figure 20.19)**. Homeobox-containing sequences have also been found in regulatory genes of much more distantly related eukaryotes, including plants and yeasts. From these similarities, we can deduce that the homeobox DNA sequence evolved very early in the history of life and was sufficiently valuable to organisms to have been conserved in animals and plants virtually unchanged for hundreds of millions of years.

Homeotic genes in animals were named *Hox* genes, short for homeobox-containing genes, because homeotic genes were the first genes found to have this sequence. Other homeobox-containing genes were later found that do not act as homeotic genes; that is, they do not directly control the identity of body parts. However, most of these genes, in animals at least, are associated with development, suggesting their ancient and fundamental importance in that process. In *Drosophila*, for example, homeoboxes are present not only in the homeotic genes but also in the egg-polarity gene *bicoid* (see Figures 18.21 and 18.22), in several of the segmentation genes, and in a master regulatory gene for eye development.

Researchers have discovered that the homeobox-encoded homeodomain binds to DNA when the protein functions as a transcription factor. Elsewhere in the protein, domains that

Figure 20.19 Conservation of homeotic genes in a fruit fly and a mouse. Homeotic genes that control the form of anterior and posterior structures of the body occur in the same linear sequence on chromosomes in *Drosophila* and mice. Each colored band on the chromosomes shown here represents a homeotic gene. In fruit flies, all homeotic genes are found on one chromosome. The mouse and other mammals have the same or similar sets of genes on four chromosomes. The color code indicates the parts of the embryos in which these genes are expressed and the adult body regions that result. All of these genes are essentially identical in flies and mice, except for those represented by black bands, which are less similar in the two animals.



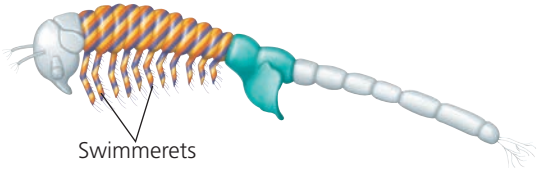
are more variable interact with other transcription factors, allowing the homeodomain-containing protein to recognize specific enhancers and regulate the associated genes. Proteins with homeodomains probably regulate development by coordinating the transcription of batteries of developmental genes, switching them on or off. In embryos of *Drosophila* and other animal species, different combinations of homeobox genes are active in different parts of the embryo. This selective expression of regulatory genes, varying over time and space, is central to pattern formation.

Developmental biologists have found that in addition to homeotic genes, many other genes involved in development are highly conserved from species to species. These include numerous genes encoding components of signaling pathways. The extraordinary similarity among some developmental genes in different animal species raises a question: How can the same genes be involved in the development of animals whose forms are so very different from each other?

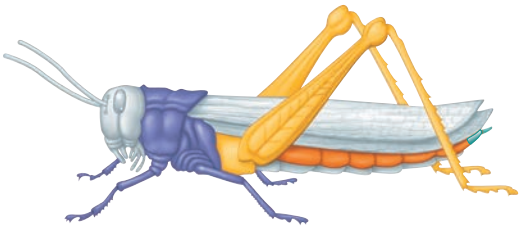
Ongoing studies are suggesting answers to this question. In some cases, small changes in regulatory sequences of particular genes cause changes in gene expression patterns that can lead to major changes in body form. For example, the differing patterns of expression of the *Hox* genes along the body axis in a crustacean and an insect can explain the variation in number of leg-bearing segments among these closely related animals **(Figure 20.20)**. In other cases, similar genes direct different developmental processes in various organisms, resulting in diverse body shapes. Several *Hox* genes, for instance, are expressed in the embryonic and larval stages of the sea urchin, a nonsegmented animal that has a body plan quite different from those of insects and mice. Sea urchin adults make the pincushion-shaped shells you may have seen on the beach; two species of live sea urchins are shown in the photo. Sea urchins are among the organisms long used in classical embryological studies (see Concept 46.2).

In this chapter of the genetics unit, you have learned how studying genomic composition and comparing the genomes of different species can illuminate the process by which genomes evolve. Furthermore, comparing developmental programs, we can see that the unity of life is reflected in the similarity of molecular and cellular mechanisms used to establish body pattern, although the genes directing development may differ among organisms. The similarities between genomes reflect the common ancestry of life on Earth. But the differences are also crucial, for they have created the huge diversity of organisms that have evolved. In the remaining chapters, we expand our perspective beyond the level of molecules, cells, and genes to explore this diversity on the organismal level.

Figure 20.20 Effect of differences in *Hox* gene expression in a crustacean and an insect. Changes in the expression patterns of *Hox* genes have occurred over evolutionary time since insects diverged from a crustacean ancestor. These changes account in part for the different body plans of **(a)** the brine shrimp *Artemia*, a crustacean, and **(b)** the grasshopper, an insect. Shown here are regions of the adult body color-coded for expression of four *Hox* genes that determine the formation of particular body parts during embryonic development. Each color represents a specific *Hox* gene.



(a) Expression of four *Hox* genes in the brine shrimp *Artemia*. Three of the *Hox* genes are expressed together in one region (indicated by stripes), specifying the identity of the segments that have swimmerets. The fourth (teal) specifies the identity of the genital segments.



(b) Expression of the grasshopper versions of the same four *Hox* genes. In the grasshopper, each *Hox* gene is expressed in a discrete region and specifies the identity of that region.

CONCEPT CHECK 20.6

1. Would you expect the genome of the macaque (a monkey) to be more like that of a mouse or that of a human? Explain.
2. DNA sequences called homeoboxes help homeotic genes in animals direct development. Given that they are common to flies and mice, explain why these animals are so different.
3. **WHAT IF? ➤** There are three times as many *Alu* elements in the human genome as in the chimpanzee genome. How do you think these extra *Alu* elements arose in the human genome? Propose a role they might have played in the divergence of these two species.

For suggested answers, see Appendix A.

20 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 20.1

VOCAB SELF-QUIZ goo.gl/Rn5Uax

The Human Genome Project fostered development of faster, less expensive sequencing techniques *(pp. 475–476)*

- The **Human Genome Project** was largely completed in 2003, aided by major advances in sequencing technology.
- In the **whole-genome shotgun** approach, the whole genome is cut into many small, overlapping fragments that are sequenced; computer software then assembles the genome sequence.

How did the Human Genome Project result in more rapid, less expensive DNA-sequencing technology?

CONCEPT 20.2

Scientists use bioinformatics to analyze genomes and their functions *(pp. 476–480)*

- Computer analysis of genome sequences aids **gene annotation**, the identification of protein-coding sequences. Methods to determine gene function include comparing sequences of newly discovered genes with those of known genes in other species and observing the effects of experimentally inactivating the genes.
- In **systems biology**, scientists use the computer-based tools of **bioinformatics** to compare genomes and study sets of genes and proteins as whole systems (**genomics** and **proteomics**). Studies include large-scale analyses of protein interactions, functional DNA elements, and genes contributing to medical conditions.

What has been the most significant finding of the ENCODE project? Why was the project expanded to include non-human species?

CONCEPT 20.3

Genomes vary in size, number of genes, and gene density *(pp. 480–482)*

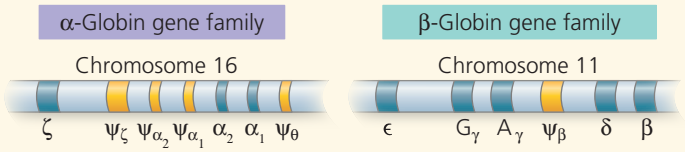
	Bacteria	Archaea	Eukarya
Genome size	Most are 1–6 Mb		Most are 10–4,000 Mb, but a few are much larger
Number of genes	1,500–7,500		Most are 5,000–45,000
Gene density	Higher than in eukaryotes		Lower than in prokaryotes (Within eukaryotes, lower density is correlated with larger genomes.)
Introns	None in protein-coding genes	Present in some genes	Present in most genes of multicellular eukaryotes, but only in some genes of unicellular eukaryotes
Other noncoding DNA	Very little		Can exist in large amounts; generally more repetitive noncoding DNA in multicellular eukaryotes

Compare genome size, gene number, and gene density (a) in the three domains and (b) among eukaryotes.

CONCEPT 20.4

Multicellular eukaryotes have a lot of noncoding DNA and many multigene families *(pp. 482–485)*

- Only 1.5% of the human genome codes for proteins or gives rise to rRNAs or tRNAs; the rest is noncoding DNA, including **pseudogenes** and **repetitive DNA** of unknown function.
- The most abundant type of repetitive DNA in multicellular eukaryotes consists of **transposable elements** and related sequences. In eukaryotes, there are two types of transposable elements: **transposons**, which move via a DNA intermediate, and **retrotransposons**, which are more prevalent and move via an RNA intermediate.
- Other repetitive DNA includes short, noncoding sequences that are tandemly repeated thousands of times (**simple sequence DNA**, which includes **STRs**); these sequences are especially prominent in centromeres and telomeres, where they probably play structural roles in the chromosome.
- Though many eukaryotic genes are present in one copy per haploid chromosome set, others (most, in some species) are members of a gene family, such as the human globin gene families:



Explain how the function of transposable elements might account for their prevalence in human noncoding DNA.

CONCEPT 20.5

Duplication, rearrangement, and mutation of DNA contribute to genome evolution *(pp. 486–491)*

- Errors in cell division can lead to extra copies of all or part of entire chromosome sets, which may then diverge if one set accumulates sequence changes. Polyploidy occurs more often among plants than animals and contributes to speciation.
- The chromosomal organization of genomes can be compared among species, providing information about evolutionary relationships. Within a given species, rearrangements of chromosomes are thought to contribute to the emergence of new species.
- The genes encoding the various related but different globin proteins evolved from one common ancestral globin gene, which duplicated and diverged into α-globin and β-globin ancestral genes. Subsequent duplication and random mutation gave rise to the present globin genes, all of which code for oxygen-binding proteins. The copies of some duplicated genes have diverged so much that the functions of their encoded proteins (such as lysozyme and α-lactalbumin) are now substantially different.
- Rearrangement of exons within and between genes during evolution has led to genes containing multiple copies of similar exons and/or several different exons derived from other genes.
- Movement of transposable elements or recombination between copies of the same element can generate new sequence combinations that are beneficial to the organism. These may alter the functions of genes or their patterns of expression and regulation.

How could chromosomal rearrangements lead to the emergence of new species?

CONCEPT 20.6

Comparing genome sequences provides clues to evolution and development *(pp. 491–496)*

- Comparisons of genomes from widely divergent and closely related species provide valuable information about ancient and more recent evolutionary history, respectively. Analysis of single nucleotide polymorphisms (SNPs) and copy-number variants (CNVs) among individuals in a species can also shed light on the evolution of that species.
- Evolutionary developmental (**evo-devo**) biologists have shown that homeotic genes and some other genes associated with animal development contain a **homeobox** region whose sequence is highly conserved among diverse species. Related sequences are present in the genes of plants and yeasts.

? *What type of information can be obtained by comparing the genomes of closely related species? Of very distantly related species?*

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–3 can be found in the Study Area in MasteringBiology.

4. DRAW IT Below are the amino acid sequences (using the single-letter code; see Figure 5.14) of four short segments of the FOXP2 protein from six species: chimpanzee (C), orangutan (O), gorilla (G), rhesus macaque (R), mouse (M), and human (H). These segments contain all of the amino acid differences between the FOXP2 proteins of these species.

PRACTICE TEST
goo.gl/iAsVgL

1. ATETI…PKSSD…TSSTT…NARRD
2. ATETI…PKSSE…TSSTT…NARRD
3. ATETI…PKSSD…TSSTT…NARRD
4. ATETI…PKSSD…TSSNT…SARRD
5. ATETI…PKSSD…TSSTT…NARRD
6. VTETI…PKSSD…TSSTT…NARRD

Use a highlighter to color any amino acid that varies among the species. (Color that amino acid in all sequences.)

(a) The C, G, R sequences are identical. Identify which lines correspond to those sequences.

(b) The H sequence differs from that of the C, G, R species at two amino acids. Underline the two differences in the H sequence.

(c) The O sequence differs from the C, G, R sequences at one amino acid (having V instead of A) and from the H sequence at three amino acids. Identify the O sequence.

(d) In the M sequence, circle the amino acid(s) that differ from the C, G, R sequences, and draw a square around those that differ from the H sequence.

(e) Primates and rodents diverged between 60 and 100 million years ago, and chimpanzees and humans about 6 million years ago. Compare the amino acid differences between the mouse and the C, G, R species with those between the human and the C, G, R species. What can you conclude?

5. EVOLUTION CONNECTION The transcription factor *FOXP2* has been found to be important for vocalization in humans. How would a comparison of the sequences of this protein across primates help us understand the evolution of speech?

6. SCIENTIFIC INQUIRY The scientists mapping the SNPs in the human genome noticed that groups of SNPs tended to be inherited together, in blocks known as haplotypes, ranging in length from 5,000 to 200,000 base pairs. There are as few as four or five commonly occurring combinations of SNPs per haplotype. Integrating what you've learned throughout this chapter and this unit, propose an explanation for this observation.

7. WRITE ABOUT A THEME: INFORMATION The continuity of life is based on heritable information in the form of DNA. In a short essay (100–150 words), explain how mutations in protein-coding genes and regulatory DNA contribute to evolution.

8. SYNTHESIZE YOUR KNOWLEDGE

Insects have three thoracic (trunk) segments. While researchers have found insect fossils with pairs of wings on all three segments, modern insects have wings or related structures on only the second and third segment. It turns out that in modern insects, *Hox* gene products act to inhibit wing formation on the first segment. The treehopper insect (above) is somewhat of an exception. In addition to having wings on its second segment, the treehopper's first segment has an ornate helmet that resembles a set of thorns, which a recent study has found to be a modified, fused pair of "wings." The thorn-like structure helps to camouflage the treehopper in tree branches, thus reducing its risk of predation. Explain how changes in gene regulation could have led to the evolution of such a structure.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

UNIT 4 EVOLUTION

A Nobel Prize laureate, Jack W. Szostak is Professor of Genetics at Harvard Medical School and Professor of Chemistry and Chemical Biology at Harvard University and is the Alexander Rich Distinguished Investigator at Massachusetts General Hospital. Dr. Szostak grew up in Canada and received a B.Sc. from McGill University and a Ph.D. in biochemistry from Cornell University. He has made pioneering contributions to genetics, including identifying the role of programmed DNA double-strand breaks in meiotic recombination and the discovery of how chromosomes are protected by telomeres. A member of the National Academy of Sciences, Dr. Szostak is now conducting groundbreaking research on the origin of life on Earth.

An Interview with Jack Szostak

You won a Nobel Prize for your work on telomeres. How did you transition from that topic to studying the origin of life?

After working on telomeres and DNA damage repair in yeast, I started to look around for something new to study. At that time, Sid Altman and Tom Cech had just discovered ribozymes, which are RNA molecules that can function as enzyme-like catalysts. I found this really exciting, and so I switched the focus of my lab to RNA biochemistry. At first, we looked at molecular evolution. But it's one thing to study evolution in the lab, where you, the experimenter, control the conditions under which an existing ribozyme can evolve. I started to wonder: How in the first place did RNA form and replicate itself on early Earth? I became more and more interested in that question, and now my whole lab is working on it.

What approaches do you use to study the origin of life?

We're using chemical experiments to try to figure out the conditions under which a primitive cell could form, grow, and divide. Much work from other laboratories has contributed to our current understanding of how the chemical building blocks of biology could have been made on early Earth. This has set the stage for exploring the next question: How did these inanimate chemicals assemble into larger structures and start acting like a living cell?

What hurdles remain to answer this question?

The first living cells must have been extremely simple. But they also must have had some of the universal features of cells today, such as a cell membrane and a genetic material like RNA. We've learned how to assemble cell-like structures with a membrane that surrounds RNA. The membranes of these structures can grow and produce "offspring" similar to themselves. The big remaining hurdle is to discover conditions under which RNA can replicate itself in a way that is compatible with a primitive cell membrane.

"All present-day life shares many features, including the same underlying biochemistry. We're trying to find explanations for such shared features."

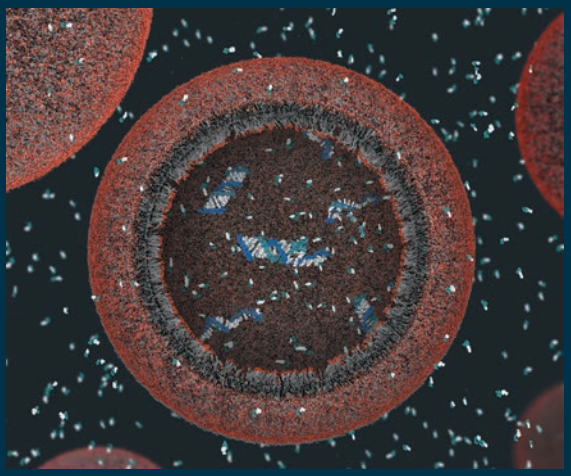
What can you learn about present-day cells by studying the origin of life?

All present-day life shares many features, including the same underlying biochemistry. We're trying to find explanations for such shared features. We've learned that the essential role of RNA seems to emerge naturally from chemical conditions like those on early Earth. But questions remain. For example, all cells maintain certain concentration gradients across their membranes, such as higher potassium levels inside the cell than out, and lower sodium levels inside than out. Cells spend a lot of energy maintaining these gradients—why is that? We don't have a good answer for that yet, but we hope to learn more about such puzzling features of life today by studying its origins.

What do you find most rewarding about your job?

I especially enjoy talking with people about new experiments. The lab is filled with talented students and post-docs who have different backgrounds and different scientific interests. Everyone has to learn how to talk with each other. This makes the lab an exciting and fun place where we all are working together to solve a big scientific problem.

Model of a protocell containing nucleotides and short bits of RNA



21 How Evolution Works

Figure 21.1 How is its resemblance to a fallen leaf helpful to this moth?

KEY CONCEPTS

21.1 The Darwinian revolution challenged traditional views of a young Earth inhabited by unchanging species

21.2 Descent with modification by natural selection explains the adaptations of organisms and the unity and diversity of life

21.3 Evolution is supported by an overwhelming amount of scientific evidence

Juvenile stage (caterpillar) of the dead-leaf moth

Endless Forms Most Beautiful

A hungry bird in the Peruvian rain forest would have to look very closely to spot a "dead-leaf moth" (*Oxytenis modestia*), which blends in well with its forest floor habitat **(Figure 21.1)**. This distinctive moth is a member of a diverse group, the more than 120,000 species of lepidopteran insects (moths and butterflies). All lepidopterans have a juvenile stage characterized by a well-developed head and many chewing mouthparts: the ravenous, efficient feeding machines we call caterpillars. (The caterpillar stage of the dead-leaf moth is also protected by its appearance: When threatened, it weaves its head back and forth, resembling a snake about to strike.) As adults, all lepidopterans share other features, such as three pairs of legs and two pairs of wings covered with small scales. But the many lepidopterans also differ from one another. How did there come to be so many different moths and butterflies, and what causes their similarities and differences?

The moth in Figure 21.1 and its many close relatives illustrate three key observations about life:

- the striking ways in which organisms are suited for life in their environments (Here and throughout this text, the term *environment* refers to other organisms as well as to the physical aspects of an organism's surroundings.)
- the many shared characteristics (unity) of life
- the rich diversity of life

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

More than a century and a half ago, Charles Darwin was inspired to develop a scientific explanation for these three broad observations. When he published his hypothesis in his book *The Origin of Species*, Darwin ushered in a scientific revolution—the era of evolutionary biology.

For now, we will define **evolution** as *descent with modification*, a phrase Darwin used in proposing that Earth's many species are descendants of ancestral species that were different from the present-day species. Evolution can also be defined as a change in the genetic composition of a population from generation to generation (see Concept 23.3).

We can also view evolution in two related but different ways: as a pattern and as a process. The *pattern* of evolutionary change is revealed by data from many scientific disciplines, including biology, geology, physics, and chemistry. These data are facts—they are observations about the natural world—and these observations show that life has evolved over time. The *process* of evolution consists of the mechanisms that cause the observed pattern of change. These mechanisms represent natural causes of the natural phenomena we observe. Indeed, the power of evolution as a unifying theory is its ability to explain and connect a vast array of observations about the living world.

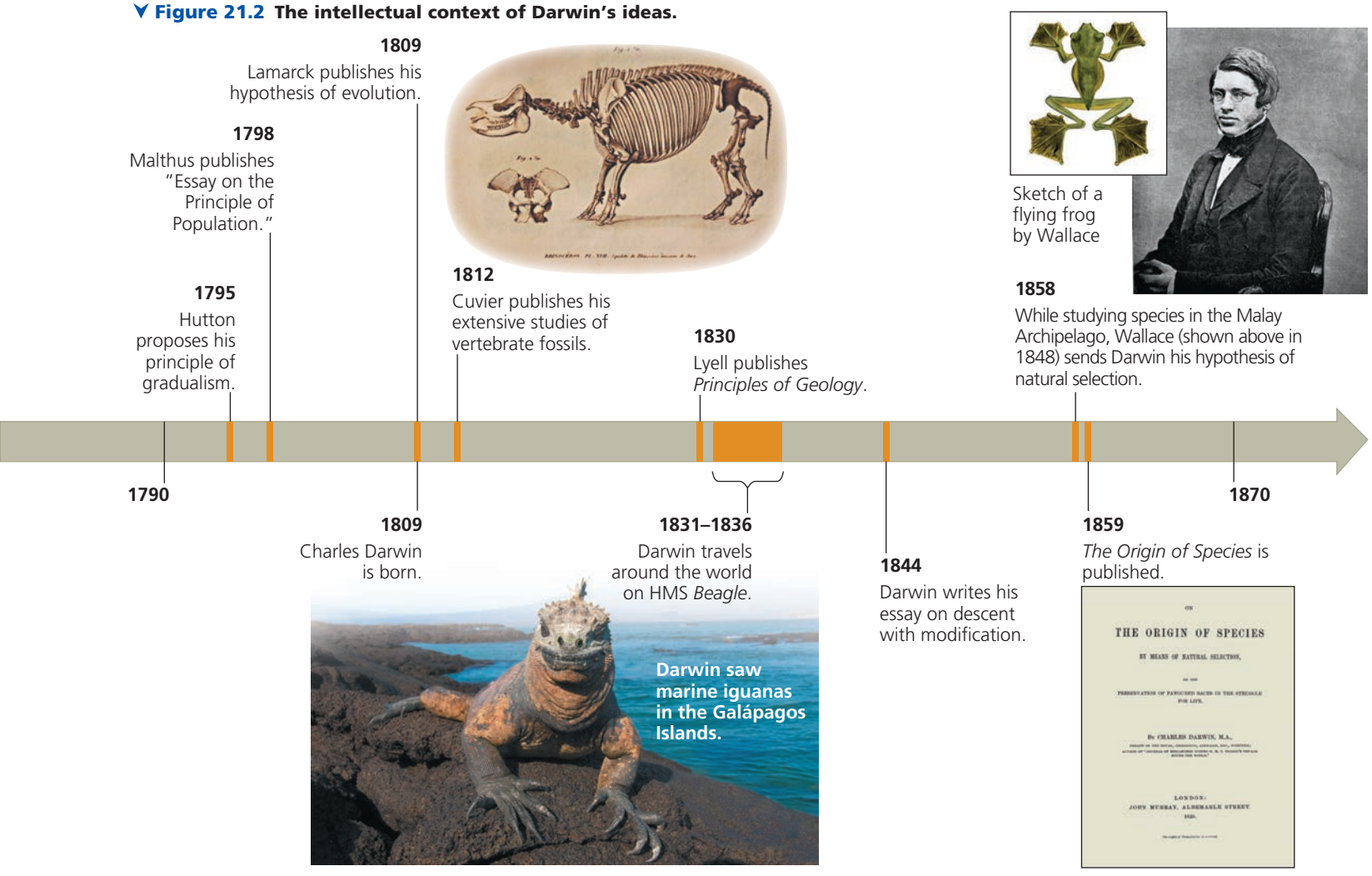
As with all general theories in science, we continue to test our understanding of evolution by examining whether it can account for new observations and experimental results. In this and the following chapters, we'll examine how ongoing discoveries shape what we know about the pattern and process of evolution. To set the stage, we'll first retrace Darwin's quest to explain the adaptations, unity, and diversity of what he called life's "endless forms most beautiful."

CONCEPT 21.1

The Darwinian revolution challenged traditional views of a young Earth inhabited by unchanging species

What impelled Darwin to challenge the prevailing views about Earth and its life? Darwin developed his revolutionary proposal over time, influenced by the work of others and by his travels **(Figure 21.2)**. As we'll see, his ideas also had deep historical roots.

Figure 21.2 The intellectual context of Darwin's ideas.



Scala Naturae and Classification of Species

Long before Darwin was born, several Greek philosophers suggested that life might have changed gradually over time. But one philosopher who greatly influenced early Western science, Aristotle (384–322 BCE), viewed species as fixed (unchanging). Through his observations of nature, Aristotle recognized certain "affinities" among organisms. He concluded that life-forms could be arranged on a ladder, or scale, of increasing complexity, later called the *scala naturae* ("scale of nature"). Each form of life, perfect and permanent, had its allotted rung on this ladder.

These ideas were generally consistent with the Old Testament account of creation, which holds that species were individually designed by God and therefore perfect. In the 1700s, many scientists interpreted the often remarkable match of organisms to their environment as evidence that the Creator had designed each species for a particular purpose.

One such scientist was Carolus Linnaeus (1707–1778), a Swedish physician and botanist who sought to classify life's diversity, in his words, "for the greater glory of God." In the 1750s, Linnaeus developed the two-part, or *binomial*, format for naming species (such as *Homo sapiens* for humans) that is still used today. In contrast to the linear hierarchy of the *scala naturae*, Linnaeus adopted a nested classification system, grouping similar species into increasingly general categories. For example, similar species are grouped in the same genus, similar genera (plural of genus) are grouped in the same family, and so on.

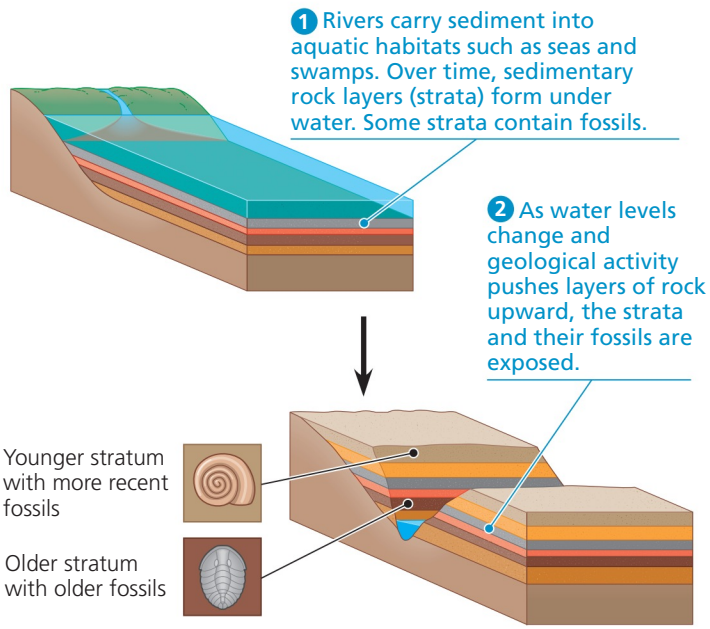
Linnaeus did not ascribe the resemblances among species to evolutionary kinship, but rather to the pattern of their creation. A century later, however, Darwin argued that classification should be based on evolutionary relationships. He also noted that scientists using the Linnaean system often grouped organisms in ways that reflected those relationships.

Ideas About Change over Time

Among other sources of information, Darwin drew from the work of scientists studying **fossils**, the remains or traces of organisms from the past. Many fossils are found in sedimentary rocks formed from the sand and mud that settle to the bottom of seas, lakes, and swamps **(Figure 21.3)**. New layers of sediment cover older ones and compress them into superimposed layers of rock called **strata** (singular, *stratum*). The fossils in a particular stratum provide a glimpse of some of the organisms that populated Earth at the time that layer formed. Later, erosion may carve through upper (younger) strata, revealing deeper (older) strata that had been buried.

Paleontology, the study of fossils, was developed in large part by French scientist Georges Cuvier (1769–1832). In examining strata near Paris, Cuvier noted that the older the stratum, the more dissimilar its fossils were to current life-forms. He also observed that from one layer to the next, some new species

Figure 21.3 Formation of sedimentary strata with fossils.



MB **Video: Grand Canyon**

appeared while others disappeared. He inferred that extinctions must have been a common occurrence, but he staunchly opposed the idea of evolution. Cuvier speculated that each boundary between strata represented a sudden catastrophic event, such as a flood, that had destroyed many of the species living in that area. Such regions, he reasoned, were later repopulated by different species immigrating from other areas.

In contrast to Cuvier's emphasis on sudden events, other scientists suggested that profound change could take place through the cumulative effect of slow but continuous processes. In 1795, Scottish geologist James Hutton (1726–1797) proposed that Earth's geologic features could be explained by gradual mechanisms, such as valleys being formed by rivers. The leading geologist of Darwin's time, Charles Lyell (1797–1875), incorporated Hutton's thinking into his proposal that the same geologic processes are operating today as in the past, and at the same rate.

Hutton's and Lyell's ideas strongly influenced Darwin's thinking. Darwin agreed that if geologic change results from slow, continuous actions rather than from sudden events, then Earth must be much older than the widely accepted age of a few thousand years. It would, for example, take a very long time for a river to carve a canyon by erosion. He later reasoned that perhaps similarly slow and subtle processes could produce substantial biological change. However, Darwin was not the first to apply the idea of gradual change to biological evolution.

Lamarck's Hypothesis of Evolution

Although some 18th-century naturalists suggested that life evolves as environments change, only one proposed a mechanism for *how* life changes over time: French biologist

Jean-Baptiste de Lamarck (1744–1829). Alas, Lamarck is primarily remembered today *not* for his visionary recognition that evolutionary change explains patterns in fossils and the match of organisms to their environments, but for the incorrect mechanism he proposed.

Lamarck published his hypothesis in 1809, the year Darwin was born. By comparing living species with fossil forms, Lamarck had found what appeared to be several lines of descent, each a chronological series of older to younger fossils leading to a living species. He explained his findings using two principles that were widely accepted at the time. The first was *use and disuse*, the idea that parts of the body that are used extensively become larger and stronger, while those that are not used deteriorate. Among many examples, he cited a giraffe stretching its neck to reach leaves on high branches. The second principle, *inheritance of acquired characteristics*, stated that an organism could pass these modifications to its offspring. Lamarck reasoned that the long, muscular neck of the living giraffe had evolved over many generations as giraffes stretched their necks ever higher.

Lamarck also thought that evolution happens because organisms have an innate drive to become more complex. Darwin rejected this idea, but he, too, thought that variation was introduced into the evolutionary process in part through inheritance of acquired characteristics. Today, however, our understanding of genetics refutes this mechanism: Experiments show that traits acquired by use during an individual's life are not inherited in the way proposed by Lamarck **(Figure 21.4)**.

➤ Figure 21.4 Acquired traits cannot be inherited. This bonsai tree was "trained" to grow as a dwarf by pruning and shaping. However, seeds from this tree would produce offspring of normal size.

Lamarck was vilified in his own time, especially by Cuvier, who denied that species ever evolve. In retrospect, however, Lamarck did recognize that the fact that organisms are well-suited for life in their environments can be explained by gradual evolutionary change, and he did propose a testable explanation for how this change occurs.

CONCEPT CHECK 21.1

1. How did Hutton's and Lyell's ideas influence Darwin's thinking about evolution?
2. **MAKE CONNECTIONS ➤** Scientific hypotheses must be testable (see Concept 1.3). Applying this criterion, are Cuvier's explanation of the fossil record and Lamarck's hypothesis of evolution scientific? Explain your answer in each case.

For suggested answers, see Appendix A.

CONCEPT 21.2

Descent with modification by natural selection explains the adaptations of organisms and the unity and diversity of life

As the 19th century dawned, it was generally thought that species had remained unchanged since their creation. A few clouds of doubt about the permanence of species were beginning to gather, but no one could have forecast the thundering storm just beyond the horizon. How did Charles Darwin become the lightning rod for a revolutionary view of life?

Darwin's Research

Charles Darwin (1809–1882) was born in Shrewsbury, in western England. Even as a boy, he had a consuming interest in nature. When he was not reading nature books, he was fishing, hunting, riding, and collecting insects. However, Darwin's father, a physician, could see no future for his son as a naturalist and sent him to medical school in Edinburgh. But Charles found medicine boring and surgery before the days of anesthesia horrifying. He quit medical school and enrolled at Cambridge University, intending to become a clergyman. (At that time, many scholars of science belonged to the clergy.)

At Cambridge, Darwin became the protégé of John Henslow, a botany professor. Soon after Darwin graduated, Henslow recommended him to Captain Robert FitzRoy, who was preparing the survey ship HMS *Beagle* for a long voyage around the world. Darwin would pay his own way and serve as a conversation partner to the young captain. FitzRoy, who was himself an accomplished scientist, accepted Darwin because he was a skilled naturalist and because they were of similar age and social class.

The Voyage of the Beagle

Darwin embarked from England on the *Beagle* in December 1831. The primary mission of the voyage was to chart poorly known stretches of the South American coastline. Darwin, however, spent most of his time on shore, observing and collecting thousands of plants and animals. He described features of organisms that made them well suited to such diverse environments as the humid jungles of Brazil, the expansive grasslands of Argentina, and the towering peaks of the Andes. He also noted that the plants and animals in temperate regions of South America more closely resembled species living in the South American tropics than species living in temperate regions of Europe. Furthermore, the fossils he found, though clearly different from living species, distinctly resembled the living organisms of South America.

Darwin also spent much time thinking about geology. Despite repeated bouts of seasickness, he read Lyell's *Principles of Geology* during the voyage. He experienced geologic change firsthand when a violent earthquake shook the coast of Chile, and he observed afterward that rocks along the coast had been thrust upward by several meters. Finding fossils of ocean organisms high in the Andes, Darwin inferred that the rocks containing the fossils must have been raised there by many similar earthquakes. These observations reinforced what he had learned from Lyell: Physical evidence did not support the traditional view that Earth was only a few thousand years old.

Darwin's interest in the species (or fossils) found in an area was further stimulated by the *Beagle*'s stop at the Galápagos, a group of volcanic islands located near the equator about 900 km west of South America **(Figure 21.5)**. Darwin was fascinated by the unusual organisms there. The birds he collected included several kinds of mockingbirds. These mockingbirds, though similar to each other, seemed to be different species. Some were unique to individual islands, while others lived on two or more adjacent islands. Furthermore, although the animals on the Galápagos resembled species living on the South American mainland, most of the Galápagos species were not known from anywhere else in the world. Darwin hypothesized that the Galápagos had been colonized by organisms that had strayed from South America and then diversified, giving rise to new species on the various islands.

Darwin's Focus on Adaptation

During the voyage of the *Beagle*, Darwin observed many examples of **adaptations**, inherited characteristics of organisms that enhance their survival and reproduction in specific environments. Later, as he reassessed his observations, he began to perceive adaptation to the environment and the origin of new species as closely related processes. Could a new species arise from an ancestral form by the gradual accumulation of adaptations to a different environment? From studies made years after Darwin's voyage, biologists have concluded that this is indeed what happened to a diverse group of finches found on the Galápagos Islands (see Figure 1.20). The finches' various beaks and behaviors are adapted to the specific foods available on their home

Figure 21.5 The voyage of HMS *Beagle* (December 1831–October 1836).

Darwin in 1840, after his return from the voyage

HMS *Beagle* in port

ABC News Video: Protecting the Galápagos Islands

Figure 21.6 Three examples of beak variation in Galápagos finches. The Galápagos Islands are home to more than a dozen species of closely related finches, some found only on a single island. A striking difference among them is their beaks, which are adapted for specific diets.

(a) Cactus-eater. The long, sharp beak of the common cactus finch (*Geospiza scandens*) helps it tear and eat cactus flowers and pulp.

(b) Insect-eater. The green warbler finch (*Certhidea olivacea*) uses its narrow, pointed beak to grasp insects.

(c) Seed-eater. The large ground finch (*Geospiza magnirostris*) has a large beak adapted for cracking seeds found on the ground.

MAKE CONNECTIONS ➤ *Review Figure 1.20. Circle the most recent common ancestor shared by the three species that eat insects. Are all of the descendants of that ancestor insect-eaters?*

islands **(Figure 21.6)**. Darwin realized that explaining such adaptations was essential to understanding evolution. His explanation of how adaptations arise centered on **natural selection**, a process in which individuals that have certain inherited traits tend to survive and reproduce at higher rates than do other individuals *because of* those traits.

By the early 1840s, Darwin had worked out the major features of his hypothesis. He set these ideas on paper in 1844, when he wrote a long essay on descent with modification and its underlying mechanism, natural selection. Yet he was still reluctant to publish his ideas, in part because he anticipated the uproar they would cause. During this time, Darwin continued to compile evidence in support of his hypothesis. By the mid-1850s, he had described his ideas to Lyell and a few others. Lyell, who was not yet convinced of evolution, nevertheless urged Darwin to publish on the subject before someone else came to the same conclusions and published first.

In June 1858, Lyell's prediction came true. Darwin received a manuscript from Alfred Russel Wallace (1823–1913), a British naturalist working in the South Pacific islands of the Malay Archipelago (see Figure 21.2). Wallace had developed a hypothesis of natural selection nearly identical to Darwin's. He asked Darwin to evaluate his paper and forward it to Lyell if it merited publication. Darwin complied, writing to Lyell: "Your words have come true with a vengeance. . . . I never saw a more striking coincidence . . . so all my originality, whatever it may amount to, will be smashed." On July 1, 1858, Lyell and a colleague presented Wallace's paper, along with extracts from Darwin's unpublished 1844 essay, to the Linnean Society of London. Darwin quickly finished his book, titled *On the Origin of Species by Means of Natural Selection* (commonly referred to as *The Origin of Species*), and published it the next year. Although Wallace had submitted his ideas for publication first, he admired Darwin and thought that Darwin had developed and tested the idea of natural selection so extensively that he should be known as its main architect.

Within a decade, Darwin's book and its proponents had convinced most scientists of the time that life's diversity is the product of evolution. Darwin succeeded where previous evolutionists had failed, mainly by presenting a plausible scientific mechanism with immaculate logic and an avalanche of supporting evidence.

HHMI Video: The Origin of Species: The Making of a Theory

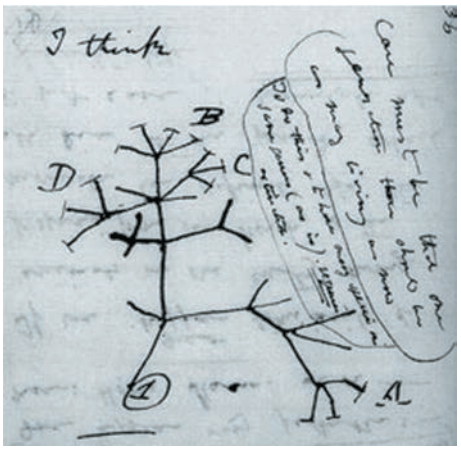
Ideas from *The Origin of Species*

In his book, Darwin amassed evidence that descent with modification by natural selection explains three broad observations about nature—the unity of life, the diversity of life, and the striking ways in which organisms are suited for life in their environments.

Descent with Modification

In the first edition of *The Origin of Species*, Darwin never used the word *evolution* (although the final word of the book is "evolved"). Rather, he discussed *descent with modification*, a phrase that summarized his view of life. Organisms share many characteristics, leading Darwin to perceive unity in life. He attributed the unity of life to the descent of all organisms from an ancestor that lived in the remote past. He also thought that as the descendants of that ancestral organism lived in various habitats, they gradually accumulated diverse modifications, or adaptations, that fit them to specific ways of life. Darwin reasoned that over a long period of time, descent with modification eventually led to the rich diversity of life we see today.

➤ Figure 21.7 "I think . . . ." In this 1837 sketch, Darwin envisioned the branching pattern of evolution. Branches that end in twigs labeled A–D represent particular groups of living organisms; all other branches represent extinct groups.



Darwin viewed the history of life as a tree, with multiple branchings from a common trunk out to the tips of the youngest twigs **(Figure 21.7)**. In his diagram, the tips of the twigs that are labeled A–D represent several groups of organisms living in the present day, while the unlabeled branches represent groups that are extinct. Each fork of the tree represents the most recent common ancestor of all the lines of evolution that subsequently branch from that point.

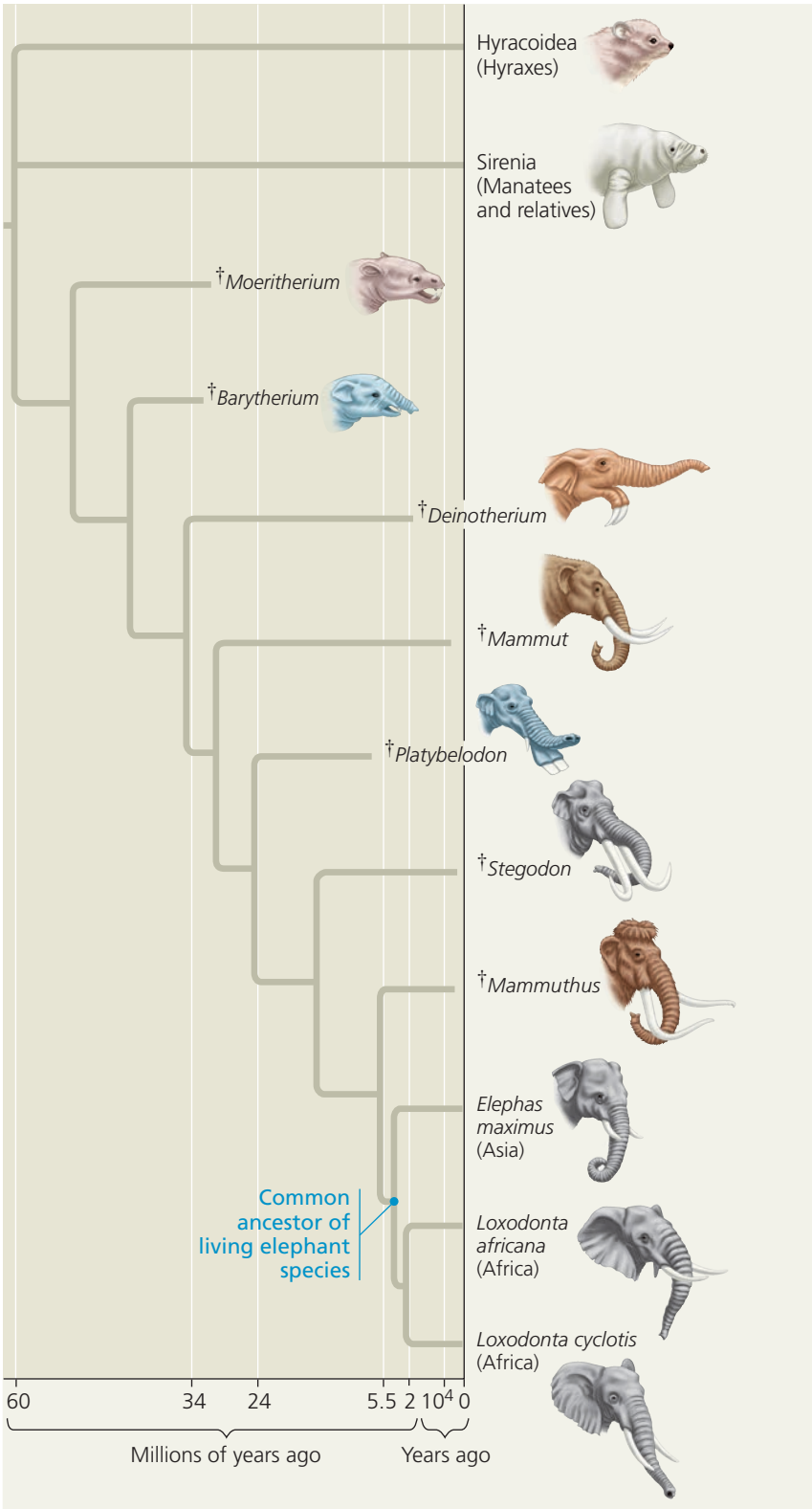
Darwin thought that such a branching process, along with past extinction events, could explain the large morphological gaps that sometimes exist between related groups of organisms. As an example, let's consider the three living species of elephants: the Asian elephant (*Elephas maximus*) and two species of African elephants (*Loxodonta africana* and *L. cyclotis*). These closely related species are very similar because they shared the same line of descent until a relatively recent split from their common ancestor, as shown in the tree diagram in **Figure 21.8**. Note that seven lineages related to elephants have become extinct over the past 32 million years. As a result, there are no living species that fill the morphological gap between the elephants and their nearest relatives today, the hyraxes and manatees.

Extinctions like those in Figure 21.8 are not uncommon. In fact, many evolutionary branches, even some major ones, are dead ends: Scientists estimate that over 99% of all species that have ever lived are now extinct. As in Figure 21.8, fossils of extinct species can document the divergence of present-day groups by "filling in" gaps between them.

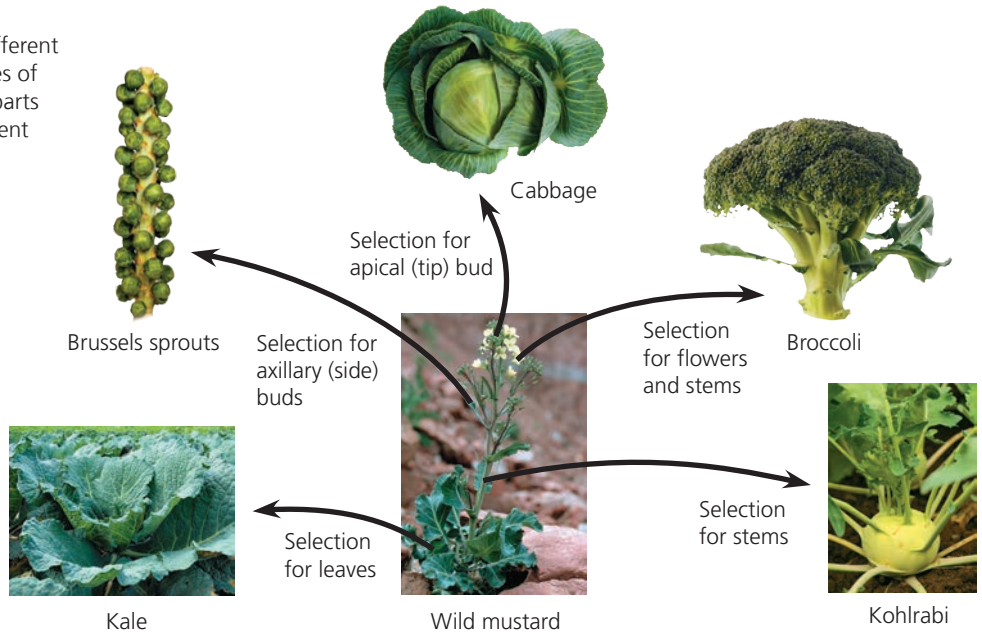
Artificial Selection, Natural Selection, and Adaptation

Darwin proposed the mechanism of natural selection to explain the observable patterns of evolution. He crafted his argument carefully, hoping to persuade even the most skeptical readers. First he discussed familiar examples of selective breeding of domesticated plants and animals. Humans have modified other species over many generations by selecting and breeding individuals that possess desired traits, a process

➤ Figure 21.8 Descent with modification. This evolutionary tree of elephants and their relatives is based mainly on fossils—their anatomy, order of appearance in strata, and geographic distribution. Note that most branches of descent ended in extinction (denoted by the dagger symbol, †). (Time line not to scale.)



VISUAL SKILLS ➤ *Based on this tree, approximately when did the most recent ancestor shared by* Mammuthus *(woolly mammoths), Asian elephants, and African elephants live?*



➤ **Figure 21.9 Artificial selection.** These different vegetables have all been selected from one species of wild mustard. By selecting variations in different parts of the plant, breeders have obtained these divergent results.

called **artificial selection** **(Figure 21.9)**. As a result of artificial selection, crops, livestock animals, and pets often bear little resemblance to their wild ancestors.

Darwin then argued that a similar process occurs in nature. He based his argument on two observations, from which he drew two inferences:

Observation #1: Members of a population often vary in their inherited traits **(Figure 21.10)**.

Observation #2: All species can produce more offspring than their environment can support **(Figure 21.11)**, and many of these offspring fail to survive and reproduce.

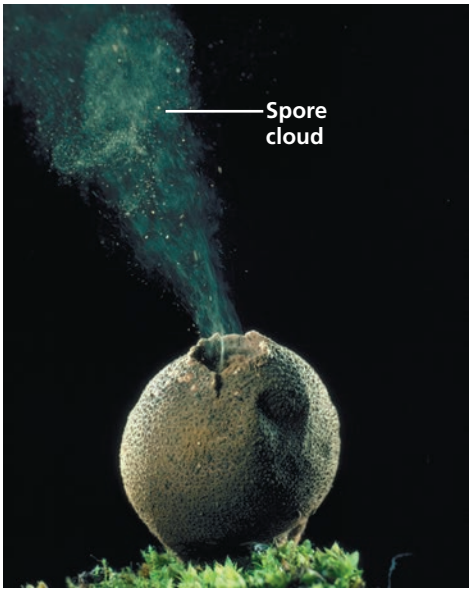
Inference #1: Individuals whose inherited traits give them a higher probability of surviving and reproducing in a given environment tend to leave more offspring than do other individuals.

Inference #2: This unequal ability of individuals to survive and reproduce will lead to the accumulation of favorable traits in the population over generations.

As these two inferences suggest, Darwin saw an important connection between natural selection and the capacity of organisms to "overreproduce." He began to make this connection after reading an essay by economist Thomas Malthus, who contended that much of human suffering—disease, famine, and war—resulted from the human population's potential to increase faster than food supplies and other resources. Similarly, Darwin realized that the capacity to overreproduce was characteristic of all species. Of the many eggs laid, young born, and seeds spread, only a tiny fraction complete their development and leave offspring of their own. The rest are eaten, starved, diseased, unmated, or unable to tolerate physical conditions of the environment such as salinity or temperature.

➤ **Figure 21.10 Variation in a population.** Individuals in this population of Asian ladybird beetles vary in color and spot pattern. Natural selection may act on these variations only if (1) they are heritable and (2) they affect the beetles' ability to survive and reproduce.

➤ **Figure 21.11 Overproduction of offspring.** A single puffball fungus can produce billions of spores that give rise to offspring. If all of these offspring and their descendants survived to maturity, they would carpet the surrounding land surface.



An organism's heritable traits can influence not only its own performance, but also how well its offspring cope with environmental challenges. For example, an organism might have a trait that gives its offspring an advantage in escaping predators, obtaining food, or tolerating physical conditions. When such advantages increase the number of offspring that survive and reproduce, the traits that are favored will likely appear at a greater frequency in the next generation. Thus, over time, natural selection resulting from factors such as predators, lack of food, or adverse physical conditions can lead to an increase in the proportion of favorable traits in a population.

How rapidly do such changes occur? Darwin reasoned that if artificial selection can bring about dramatic change in a relatively short period of time, then natural selection should be capable of substantial modification of species over many hundreds of generations. Even if the advantages of some heritable traits over others are slight, the advantageous variations will gradually accumulate in the population, and less favorable variations will diminish. Over time, this process will increase the frequency of individuals with favorable adaptations, hence increasing the degree to which organisms are well suited for life in their environment.

Key Features of Natural Selection

Let's now recap the main ideas of natural selection:

- Natural selection is a process in which individuals that have certain heritable traits survive and reproduce at a higher rate than do other individuals because of those traits.
- Over time, natural selection can increase the frequency of adaptations that are favorable in a given environment **(Figure 21.12)**.
- If an environment changes, or if individuals move to a new environment, natural selection may result in adaptation to these new conditions, sometimes giving rise to new species.

One subtle but important point is that although natural selection occurs through interactions between individual organisms and their environment, *individuals do not evolve*. Rather, it is the population that evolves over time.

A second key point is that natural selection can amplify or diminish only those heritable traits that differ among the individuals in a population. Thus, even if a trait is heritable, if all the individuals in a population are genetically identical for that trait, evolution by natural selection cannot occur.

Third, remember that environmental factors vary from place to place and over time. A trait that is favorable in one place or time may be useless—or even detrimental—in other places or times. Natural selection is always operating, but which traits are favored depends on the context in which a species lives and mates.

Figure 21.12 Camouflage as an example of evolutionary adaptation. Related species of the insects called mantises have diverse shapes and colors that evolved in different environments, as seen in this South African flower-eyed mantis (*Pseudocreobotra wahlbergi*; top) and Malaysian orchid mantis (*Hymenopus coronatus*; bottom).

VISUAL SKILLS ➤ *Use evidence from these two images to explain how these mantises demonstrate the three key observations about life introduced at the beginning of this chapter: the unity and diversity of life and the match between organisms and their environments.*

Next, we'll survey the wide range of observations that support a Darwinian view of evolution by natural selection.

CONCEPT CHECK 21.2

1. How does the concept of descent with modification explain both the unity and diversity of life?
2. **WHAT IF?** ➤ If you discovered a fossil of an extinct mammal that lived high in the Andes, would you predict that it would more closely resemble present-day mammals from South American jungles or present-day mammals that live high in Asian mountains? Explain.
3. **MAKE CONNECTIONS** ➤ Review the relationship between genotype and phenotype (see Figures 14.5 and 14.6). Suppose that in a particular pea population, flowers with the white phenotype are favored by natural selection. Predict what would happen over time to the frequency of the *p* allele in the population, and explain your reasoning.

For suggested answers, see Appendix A.

CONCEPT 21.3

Evolution is supported by an overwhelming amount of scientific evidence

In *The Origin of Species*, Darwin marshaled a broad range of evidence to support the concept of descent with modification. Still—as he readily acknowledged—there were instances in which key evidence was lacking. For example, Darwin referred to the origin of flowering plants as an "abominable mystery," and he lamented the lack of fossils showing how earlier groups of organisms gave rise to new groups.

In the last 150 years, new discoveries have filled many of the gaps that Darwin identified. The origin of flowering plants, for example, is much better understood (see Concept 30.3), and many fossils have been discovered that signify the origin of new groups of organisms (see Concept 25.2). In this section, we'll consider four types of data that document the pattern of evolution and illuminate how it occurs: direct observations, homology, the fossil record, and biogeography.

Direct Observations of Evolutionary Change

Biologists have documented evolutionary change in thousands of scientific studies. We'll examine many such studies throughout this unit, but let's look at two examples here.

Natural Selection in Response to Introduced Species

Animals that eat plants, called herbivores, often have adaptations that help them feed efficiently on their primary food sources. What happens when herbivores switch to a new food source with different characteristics?

An opportunity to study this question in nature is provided by soapberry bugs, which use their "beak"—a hollow, needle-like mouthpart—to feed on seeds located within the fruits of various plants. In southern Florida, the soapberry bug (*Jadera haematoloma*) feeds on the seeds of a native plant, the balloon vine (*Cardiospermum corindum*). In central Florida, however, balloon vines have become rare. Instead, soapberry bugs in that region now feed on the seeds of the goldenrain tree (*Koelreuteria elegans*), a species recently introduced from Asia.

Soapberry bugs feed most effectively when the length of their beak is similar to the depth at which seeds are found within the fruit. Goldenrain tree fruit consists of three flat lobes, and its seeds are much closer to the fruit surface than are the seeds of the plump, round fruit of the native balloon vine. These differences led researchers to predict that in populations that feed on goldenrain tree, natural selection would result in beaks that are *shorter* than those in populations that feed on balloon vine **(Figure 21.13)**. Indeed, beak lengths are shorter in the populations that feed on goldenrain tree.

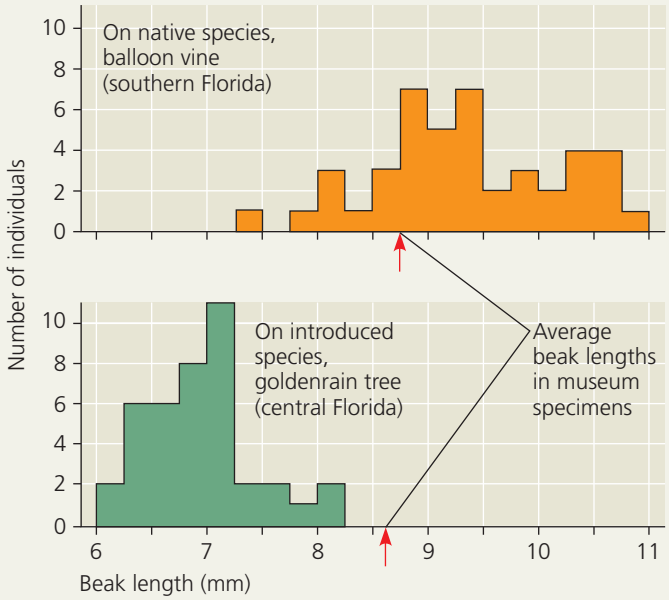
Figure 21.13

Inquiry **Can a change in a population's food source result in evolution by natural selection?**

Field Study Soapberry bugs feed most effectively when the length of their "beak" is similar to the depth of the seeds within the fruit. Scott Carroll and his colleagues measured beak lengths in soapberry bug populations feeding on the native balloon vine. They also measured beak lengths in populations feeding on the introduced goldenrain tree. The researchers then compared the measurements with those of museum specimens collected in the two areas before the goldenrain tree was introduced.

Soapberry bug with beak inserted in balloon vine fruit

Results Beak lengths were shorter in populations feeding on the introduced species than in populations feeding on the native species, in which the seeds are buried more deeply. The average beak length in museum specimens from each population (indicated by red arrows) was similar to beak lengths in populations feeding on native species.



Conclusion Museum specimens and contemporary data suggest that a change in the size of the soapberry bug's food source can result in evolution by natural selection for a corresponding change in beak size.

Data from S. P. Carroll and C. Boyd, Host race radiation in the soapberry bug: natural history with the history, *Evolution* 46:1052–1069 (1992).

WHAT IF? ➤ *Data from additional studies showed that when soapberry bug eggs from a population that fed on balloon vine fruits were then reared on goldenrain tree fruits (or vice versa), the beak lengths of the adult insects were most similar to those in the population from which the eggs were initially obtained. Interpret these results.*

Researchers have also studied beak length evolution in soapberry bug populations that feed on plants introduced to Louisiana, Oklahoma, and Australia. In each of these locations, the fruit of the introduced plants is larger than the fruit of the native plant. Thus, in populations feeding on introduced species in these regions, researchers predicted that natural selection would result in the evolution of *longer* beaks. Again, data collected in field studies upheld this prediction.

The observed changes in beak lengths had important consequences: In Australia, for example, the increase in beak length nearly doubled the success with which soapberry bugs could eat the seeds of the introduced species. Furthermore, since historical data show that the goldenrain tree reached central Florida just 35 years before the scientific studies were initiated, the results demonstrate that natural selection can cause rapid evolution in a wild population.

The Evolution of Drug-Resistant Bacteria

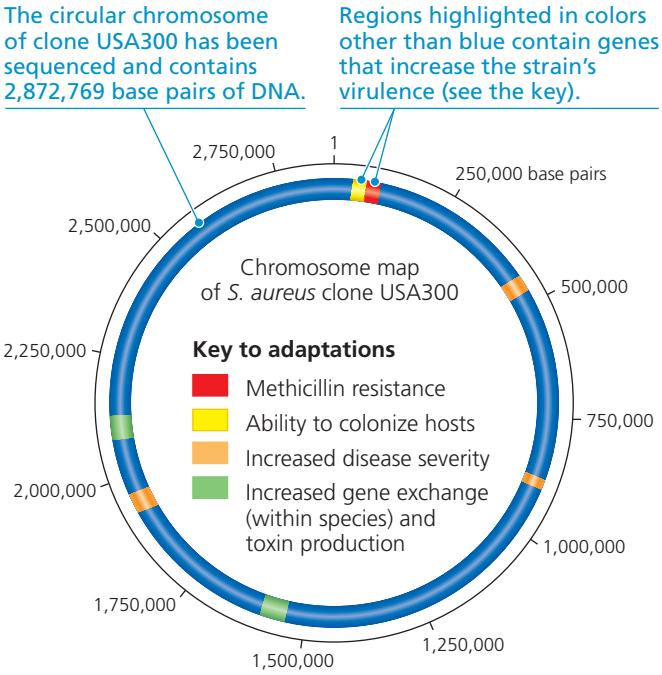
An example of ongoing natural selection that dramatically affects humans is the evolution of drug-resistant pathogens (disease-causing organisms and viruses). This is a particular problem with bacteria and viruses because they can produce new generations in a short period of time; as a result, resistant strains of these pathogens can proliferate very quickly.

Consider the evolution of drug resistance in the bacterium *Staphylococcus aureus*. About one in three people harbor this species on their skin or in their nasal passages with no negative effects. However, certain genetic varieties (strains) of this species, known as methicillin-resistant *S. aureus* (MRSA), are formidable pathogens. Most MRSA infections are caused by recently appearing strains such as clone USA300, which can cause "flesh-eating disease" and potentially fatal infections **(Figure 21.14)**. How did clone USA300 and other strains of MRSA become so dangerous?

The story begins in 1943, when penicillin became the first widely used antibiotic. Since then, penicillin and other antibiotics have saved millions of lives. However, by 1945, more than 20% of the *S. aureus* strains seen in hospitals were already resistant to penicillin. These bacteria had an enzyme, penicillinase, that could destroy penicillin. Researchers responded by developing antibiotics that were not destroyed by penicillinase, but resistance to each new drug was observed in some *S. aureus* populations within a few years.

Then, in 1959, doctors began using a promising new antibiotic, methicillin. But within two years, methicillin-resistant strains of *S. aureus* were observed. How did these resistant strains emerge? Methicillin works by deactivating an enzyme that bacteria use to synthesize their cell walls. However, some *S. aureus* populations included individuals that were able to synthesize their cell walls using a different enzyme that was not affected by methicillin. These individuals survived the methicillin treatments and reproduced at higher rates than did other individuals. Over time, these resistant individuals became increasingly common, leading to the spread of MRSA.

Figure 21.14 Clone USA300: a virulent strain of methicillin-resistant *Staphylococcus aureus* (MRSA). Resistant to multiple antibiotics and highly contagious, this strain and its close relatives can cause lethal infections of the skin, lungs, and blood. As shown here, researchers have identified key areas of the USA300 genome that code for adaptations that cause its virulent properties.



WHAT IF? ➤ *Some drugs being developed specifically target and kill only* S. aureus; *others slow the growth of MRSA but do not kill it. Based on how natural selection works and on the fact that bacterial species can exchange genes, explain why each of these strategies might be effective.*

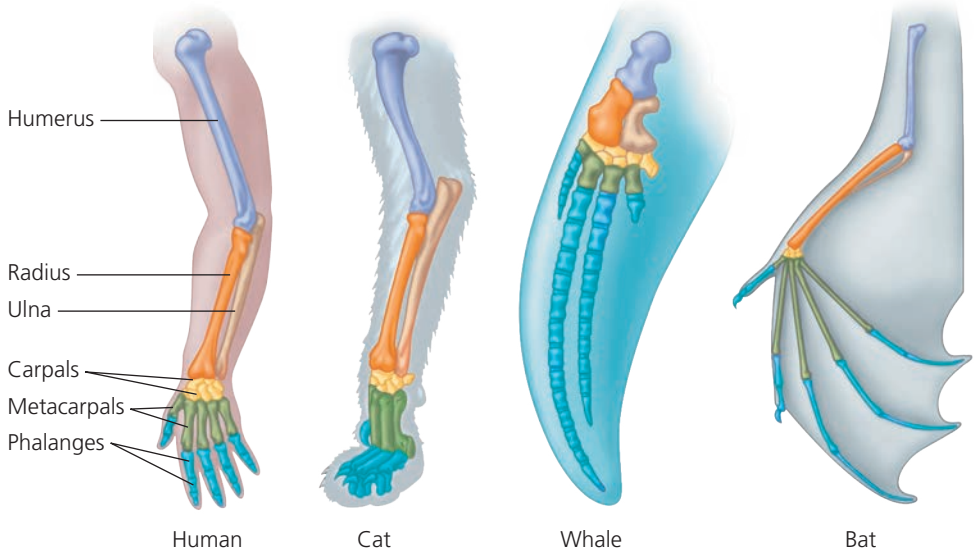
Initially, MRSA could be controlled by antibiotics that work differently from the way methicillin works. But this has become less effective because some MRSA strains are resistant to multiple antibiotics—probably because bacteria can exchange genes with members of their own and other species. Thus, the multidrug-resistant strains of today may have emerged over time as MRSA strains that were resistant to different antibiotics exchanged genes.

Finally, it is important to note that *S. aureus* is not the only pathogenic bacterium that has evolved resistance to multiple antibiotics. Furthermore, in recent decades, antibiotic resistance has spread much faster than new antibiotics have been discovered—a problem of great public health concern. Hope may loom on the horizon, however. For example, in 2015, scientists reported the discovery of "teixobactin," a new antibiotic that shows promise for treating MRSA and other pathogens. In addition, as we'll describe in the Scientific Skills Exercise in Chapter 27, the methods used in the discovery of teixobactin may lead to the discovery of other new antibiotics as well.

The *S. aureus* and soapberry bug examples highlight three key points about natural selection. First, natural selection is a process of editing, not a creative mechanism. A drug does not *create* resistant pathogens; it *selects for* resistant individuals that are already present in the population. Second, in species that produce new generations in short periods of time, evolution

by natural selection can occur rapidly—in just a few years (*S. aureus*) or decades (soapberry bugs). Third, natural selection depends on time and place. It favors those characteristics in a genetically variable population that provide an advantage in the current, local environment. What is beneficial in one situation may be useless or even harmful in another. Beak lengths suitable for the size of the typical fruit eaten by members of a particular soapberry bug population are favored by natural selection. However, a beak length suitable for fruit of one size can be disadvantageous when the bug is feeding on fruit of another size.

Figure 21.15 Mammalian forelimbs: homologous structures. Even though they have become adapted for different functions, the forelimbs of all mammals are constructed from the same basic skeletal elements: one large bone (purple), attached to two smaller bones (orange and tan), attached to several small bones (gold), attached to several metacarpals (green), attached to approximately five digits, each of which is composed of multiple phalanges (blue).



Homology

A second type of evidence for evolution comes from analyzing similarities among different organisms. As we've discussed, evolution is a process of descent with modification: Characteristics present in an ancestral organism are altered (by natural selection) in its descendants over time as they face different environmental conditions. As a result, related species can have characteristics that have an underlying similarity yet function differently. Similarity resulting from common ancestry is known as **homology**. As we'll describe in this section, an understanding of homology can be used to make testable predictions and explain observations that are otherwise puzzling.

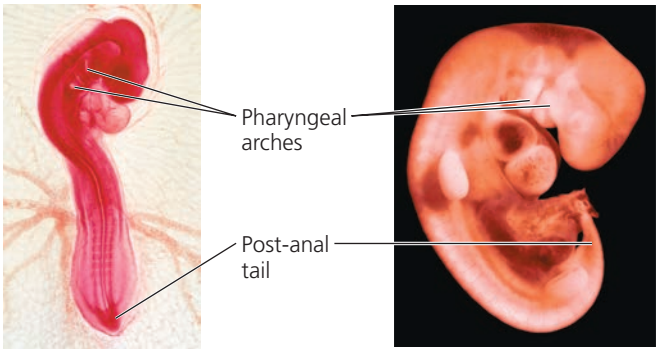
Anatomical and Molecular Homologies

The view of evolution as a remodeling process leads to the prediction that closely related species should share similar features—and they do. Of course, closely related species share the features used to determine their relationship, but they also share many other features. Some of these shared features make little sense except in the context of evolution. For example, the forelimbs of all mammals, including humans, cats, whales, and bats, show the same arrangement of bones from the shoulder to the tips of the digits, even though the appendages have very different functions: lifting, walking, swimming, and flying **(Figure 21.15)**. Such striking anatomical resemblances would be highly unlikely if these structures had arisen anew in each species. Rather, the underlying skeletons of the arms, forelegs, flippers, and wings of different mammals are **homologous structures** that represent variations on a structural theme that was present in their common ancestor.

Comparing early stages of development in different animal species reveals additional anatomical homologies not visible in adult organisms. For example, at some point in their development, all vertebrate embryos have a tail located posterior to (behind) the anus, as well as structures called pharyngeal (throat) arches **(Figure 21.16)**. These homologous throat arches ultimately develop into structures with very different functions, such as gills in fishes and parts of the ears and throat in humans and other mammals.

Some of the most intriguing homologies concern "leftover" structures of marginal, if any, importance to the organism. These **vestigial structures** are remnants of features that served a function in the organism's ancestors. For instance, the skeletons of some snakes retain vestiges of the pelvis and leg bones of walking ancestors. Another example is provided by eye remnants that are buried under scales in blind species

Figure 21.16 Anatomical similarities in vertebrate embryos. At some stage in their embryonic development, all vertebrates have a tail located posterior to the anus (referred to as a post-anal tail), as well as pharyngeal (throat) arches. Descent from a common ancestor can explain such similarities.



of cave fishes. We would not expect to see these vestigial structures if snakes and blind cave fishes had origins separate from those of other vertebrate animals.

Biologists also observe similarities among organisms at the molecular level. All forms of life use essentially the same genetic code, suggesting that all species descended from common ancestors that used this code. But molecular homologies go beyond a shared code. For example, organisms as dissimilar as humans and bacteria share genes inherited from a very distant common ancestor. Some of these homologous genes have acquired new functions, while others, such as those coding for the ribosomal subunits used in protein synthesis (see Figure 17.18), have retained their original functions. It is also common for organisms to have genes that have lost their function, even though the homologous genes in related species may be fully functional. Like vestigial structures, it appears that such inactive "pseudogenes" may be present simply because a common ancestor had them.

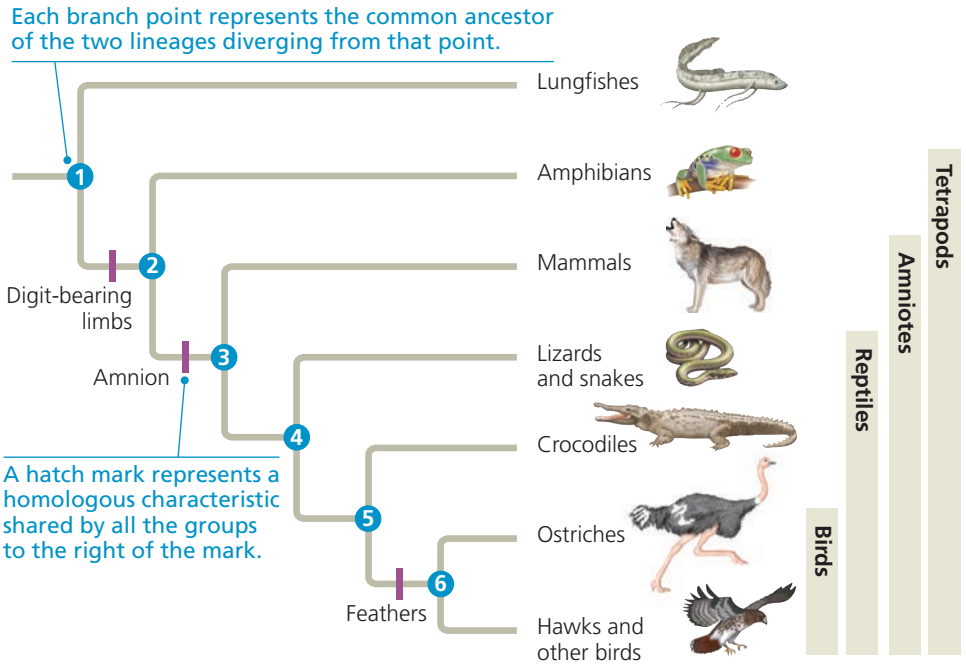
Homologies and "Tree Thinking"

Some homologous characteristics, such as the genetic code, are shared by all species because they date to the deep ancestral past. In contrast, homologous characteristics that evolved more recently are shared only within smaller groups of organisms. Consider the *tetrapods* (from the Greek *tetra*, four, and *pod*, foot), the vertebrate group that consists of amphibians, mammals, and reptiles. Like all vertebrates, tetrapods have a backbone. But unlike other vertebrates, tetrapods also have limbs with digits (see Figure 21.15). As suggested by this example, homologous characteristics form a nested pattern: All life shares the deepest layer (in this case, all vertebrates have a backbone), and each successive smaller group adds its own homologies to those it shares with larger groups (in this case, all tetrapods have a backbone *and* limbs with digits). This nested pattern is exactly what we would expect to result from descent with modification from a common ancestor.

Biologists often represent the pattern of descent from common ancestors with an **evolutionary tree**, a diagram that reflects evolutionary relationships among groups of organisms. We will explore evolutionary trees in more detail in Chapter 22, but for now, let's consider how we can interpret and use such trees.

Figure 21.17 is an evolutionary tree of tetrapods and their closest living relatives, the lungfishes. In this diagram, each branch point represents the most recent common ancestor of the two lineages diverging from that point. For example, lungfishes and all tetrapods descended from ancestor 1, whereas mammals, lizards and snakes, crocodiles, and birds all descended from ancestor 3. As expected, the three homologies shown on the tree—limbs with digits, the amnion (a protective embryonic membrane), and feathers—form a nested pattern. Limbs with digits were present in common ancestor 2 and hence are found in all of the descendants of that ancestor (the tetrapods). The amnion was present only in ancestor 3 and hence is shared only by some tetrapods (mammals and reptiles). Feathers were present only in ancestor 6 and hence are found only in birds.

To explore "tree thinking" further, note that in Figure 21.17, mammals are positioned closer to amphibians than to birds. As a result, you might conclude that mammals are more closely related to amphibians than they are to birds. However, mammals are actually more closely related to birds than to amphibians because mammals and birds share a more recent common ancestor (ancestor 3) than do mammals and amphibians (ancestor 2). Ancestor 2 is also the most recent common ancestor of birds and amphibians, making mammals and birds equally related to amphibians. Finally, note that the tree in Figure 21.17 shows the relative timing of events but not their actual dates. Thus, we can conclude that ancestor 2 lived before ancestor 3, but we do not know when that was.



➤ Figure 21.17 Tree thinking: information provided in an evolutionary tree. This evolutionary tree for tetrapods and their closest living relatives, the lungfishes, is based on anatomical and DNA sequence data. The purple bars indicate the origin of three important homologies, each of which evolved only once. Birds are nested within and evolved from reptiles; hence, the group of organisms called "reptiles" technically includes birds.

VISUAL SKILLS ➤ *Based on this evolutionary tree, are crocodiles more closely related to lizards or birds? Explain.*

MB **Figure Walkthrough**

Evolutionary trees are hypotheses that summarize our current understanding of patterns of descent. Our confidence in these relationships, as with any hypothesis, depends on the strength of the supporting data. In the case of Figure 21.17, the tree is supported by many different data sets, including both anatomical and DNA sequence data. As a result, biologists are confident that it accurately reflects evolutionary history. Scientists can use such well-supported evolutionary trees to make specific and sometimes surprising predictions about organisms (see Figure 22.17).

A Different Cause of Resemblance: Convergent Evolution

Although organisms that are closely related share characteristics because of common descent, distantly related organisms can resemble one another for a different reason: **convergent evolution**, the independent evolution of similar features in different lineages. Consider marsupial mammals, many of which live in Australia. Marsupials are distinct from another group of mammals—the placental mammals, or eutherians—few of which live in Australia. (Eutherians complete their embryonic development in the uterus, whereas marsupials are born as embryos and complete their development in an external pouch.) Some Australian marsupials have eutherian look-alikes with superficially similar adaptations. For instance, a forest-dwelling Australian marsupial called the sugar glider is superficially very similar to flying squirrels, gliding eutherians that live in North American forests **(Figure 21.18)**. But the sugar glider has many other characteristics that make it a marsupial, much more closely related to kangaroos and other Australian marsupials than to flying squirrels or other eutherians. Once again, our understanding of evolution can explain these observations. Although they evolved independently from different ancestors, these two mammals have adapted to similar environments in similar ways. In such examples in which species share features because of convergent evolution, the resemblance is said to be **analogous**, not homologous. Analogous features share similar function, but not common ancestry, while homologous features share common ancestry, but not necessarily similar function.

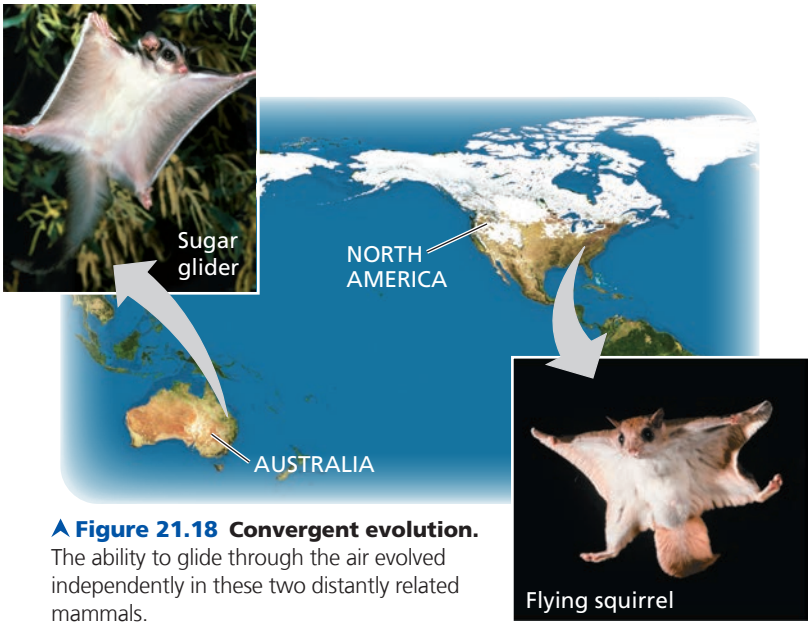


Figure 21.18 Convergent evolution. The ability to glide through the air evolved independently in these two distantly related mammals.

The Fossil Record

A third type of evidence for evolution comes from fossils. The fossil record documents the pattern of evolution, showing that past organisms differed from present-day organisms and that many species have become extinct. Fossils also show the evolutionary changes that have occurred in various groups of organisms. To give one of hundreds of possible examples, researchers found that over several thousand years, the pelvic bone in fossil stickleback fish became greatly reduced in size. The consistent nature of this change over time suggests that the reduction in the size of the pelvic bone may have been driven by natural selection.

Fossils can also shed light on the origins of new groups of organisms. An example is the fossil record of cetaceans, the mammalian order that includes whales, dolphins, and porpoises. Some of these fossils **(Figure 21.19)** provided strong support for a hypothesis based on DNA sequence data: that cetaceans are closely related to even-toed ungulates, a group that includes hippopotamuses, pigs, deer, and cows.

What else can fossils tell us about cetacean origins? The earliest cetaceans lived 50–60 million years ago. The fossil record indicates that prior to that time, most mammals were terrestrial. Although scientists had long realized that whales and other cetaceans originated from land mammals, few fossils

Figure 21.19 Ankle bones: one piece of the puzzle. Comparing fossils and present-day examples of the astragalus (a type of ankle bone) indicates that cetaceans are closely related to even-toed ungulates. **(a)** In most mammals, the astragalus is shaped like that of a dog, with a double hump on one end (red arrows) but not at the opposite end (blue arrow). **(b)** Fossils show that the early cetacean *Pakicetus* had an astragalus with double humps at both ends, a shape otherwise found only in pigs **(c)**, deer **(d)**, and all other even-toed ungulates.

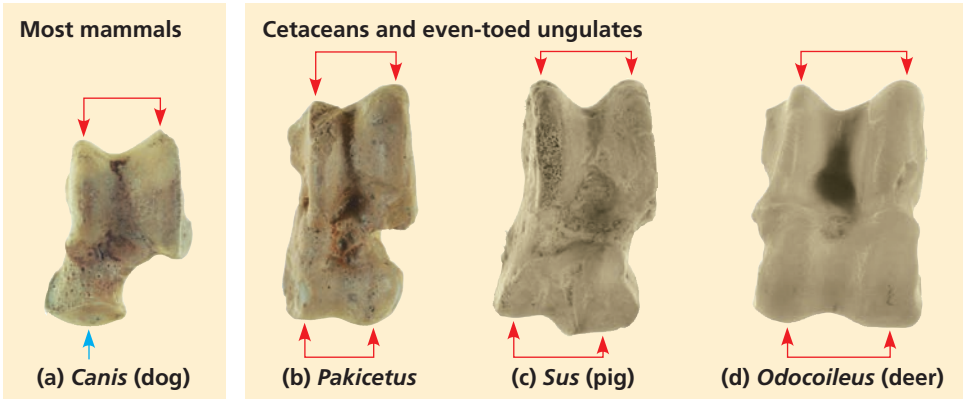
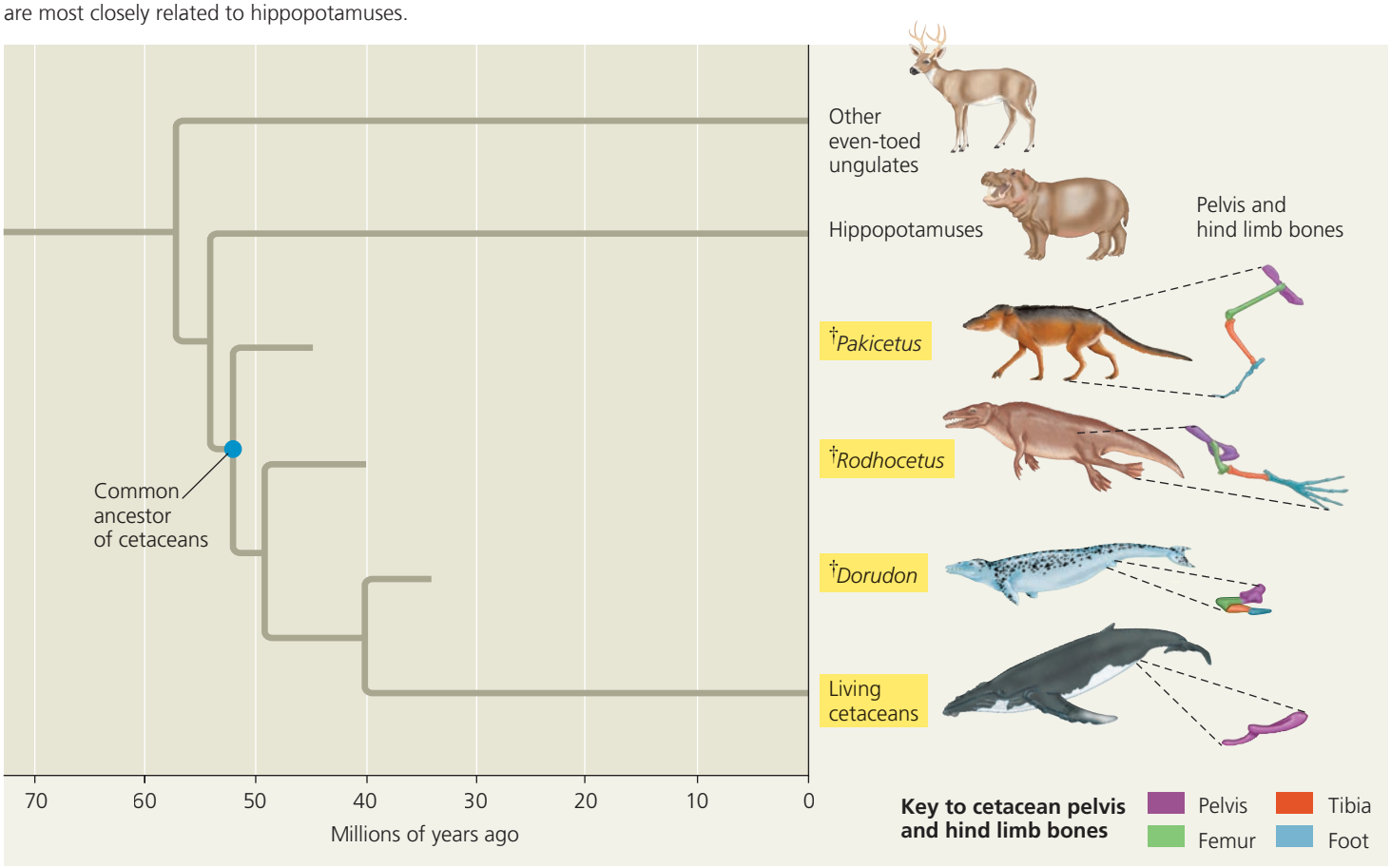


Figure 21.20 The transition to life in the sea. Multiple lines of evidence support the hypothesis that cetaceans (highlighted in yellow) evolved from terrestrial mammals. Fossils document the reduction over time in the pelvis and hind limb bones of extinct (†) cetacean ancestors, including *Pakicetus, Rodhocetus,* and *Dorudon*. DNA sequence data support the hypothesis that cetaceans are most closely related to hippopotamuses.

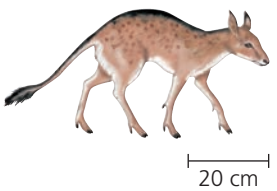


VISUAL SKILLS ➤ *Use the diagram to determine which happened first during the evolution of cetaceans: changes in hind limb structure or the origin of tail flukes. Explain.*

had been found that revealed how cetacean limb structure had changed over time, leading eventually to the loss of hind limbs and the development of flukes (the lobes on a whale's tail) and flippers. In the past few decades, however, a series of remarkable fossils have been discovered in Pakistan, Egypt, and North America. These fossils document steps in the transition from life on land to life in the sea, filling in some of the gaps between ancestral and living cetaceans **(Figure 21.20)**.

Collectively, the recent fossil discoveries document the origin of a group of mammals, the cetaceans. These discoveries also show that cetaceans and their close living relatives (hippopotamuses and other even-toed ungulates) are much more different from each other than were *Pakicetus* and early even-toed ungulates, such as *Diacodexis* **(Figure 21.21)**. Similar patterns are seen in fossils documenting the origins of other groups of organisms, including mammals (see Figure 25.7), flowering plants (see Concept 30.3), and tetrapods (see Figure 34.21).

Figure 21.21 *Diacodexis*, an early even-toed ungulate.



In each of these cases, the fossil record shows that over time, descent with modification produced increasingly large differences among related groups of organisms, ultimately resulting in the diversity of life we see today.

Biogeography

A fourth type of evidence for evolution comes from the field of **biogeography**, the scientific study of the geographic distributions of species. The geographic distributions of organisms are influenced by many factors, including *continental drift*, the slow movement of Earth's continents over time. About 250 million years ago, these movements united all of Earth's landmasses into a single large continent called **Pangaea** (see Figure 25.16). Roughly 200 million years ago, Pangaea began to break apart; by 20 million years ago, the continents we know today were within a few hundred kilometers of their present locations.

We can use our understanding of evolution and continental drift to predict where fossils of different groups of organisms might be found. For example, scientists have constructed evolutionary trees for horses based on anatomical data. These trees and the ages of fossils of horse ancestors

suggest that the genus that includes present-day horses (*Equus*) originated 5 million years ago in North America. Geologic evidence indicates that at that time, North and South America were not yet connected, making it difficult for horses to travel between them. Thus, we would predict that the oldest *Equus* fossils should be found only on the continent on which the group originated—North America. This prediction and others like it for different groups of organisms have been upheld, providing more evidence for evolution.

We can also use our understanding of evolution to explain biogeographic data. For example, islands generally have many plant and animal species that are **endemic** (found nowhere else in the world). Yet, as Darwin described in *The Origin of Species*, most island species are closely related to species from the nearest mainland or a neighboring island. He explained this observation by suggesting that islands are colonized by species from the nearest mainland. These colonists eventually give rise to new species as they adapt to their new environments. Such a process also explains why two islands with similar environments in distant parts of the world tend to be populated not by species that are closely related to each other, but rather by species related to those of the nearest mainland, where the environment is often quite different.

What Is Theoretical About Darwin's View of Life?

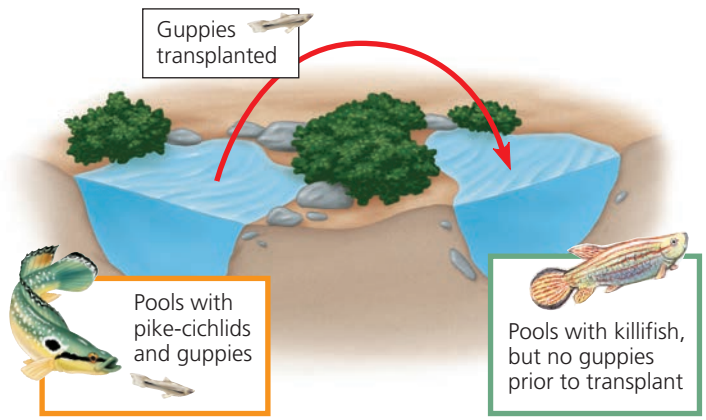
Some people dismiss Darwin's ideas as "just a theory." However, as we have seen, the *pattern* of evolution—the observation that life has evolved over time—has been documented directly and is supported by a great deal of evidence. In addition, Darwin's explanation of the *process* of evolution—that natural selection is the primary cause of the observed pattern of evolutionary change—makes sense of massive amounts of data. The effects of natural selection also can be observed and tested in nature. One such experiment is described in the **Scientific Skills Exercise**.

SCIENTIFIC SKILLS EXERCISE

Making and Testing Predictions

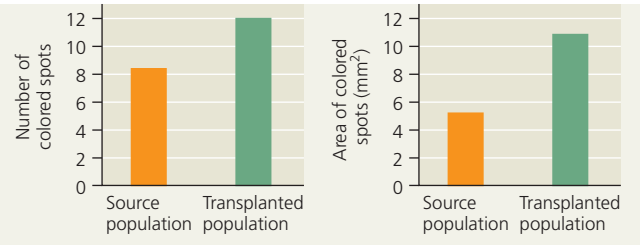
Can Predation Result in Natural Selection for Color Patterns in Guppies? Our understanding of evolution changes constantly as new observations lead to new hypotheses—and hence to new ways to test our understanding of evolutionary theory. Consider the wild guppies (*Poecilia reticulata*) that live in pools connected by streams on the Caribbean island of Trinidad. Male guppies have highly varied color patterns that are controlled by genes that are only expressed in adult males. Female guppies choose males with bright color patterns as mates more often than they choose males with drab coloring. But the bright colors that attract females also can make the males more conspicuous to predators. Researchers observed that in pools with few predator species, the benefits of bright colors appear to "win out," and males are more brightly colored than in pools where predation is more intense.

One guppy predator, the killifish, preys on juvenile guppies that have not yet displayed their adult coloration. Researchers predicted that if adult guppies with drab colors were transferred to a pool with only killifish, eventually the descendants of these guppies would be more brightly colored (because of the female preference for brightly colored males).



How the Experiment Was Done Researchers transplanted 200 guppies from pools containing pike-cichlid fish, intense predators of adult guppies, to pools containing killifish, less active predators that prey mainly on juvenile guppies. They tracked the number of bright-colored spots and the total area of those spots on male guppies in each generation.

Data from the Experiment After 22 months (15 generations), researchers compared the color pattern data for guppies from the source and transplanted populations.



Data from J. A. Endler, Natural selection on color patterns in *Poecilia reticulata*, *Evolution* 34:76–91 (1980).

INTERPRET THE DATA

1. Identify the following elements of hypothesis-based science in this example: (a) question, (b) hypothesis, (c) prediction, (d) control group, and (e) experimental group. (For additional information about hypothesis-based science, see Chapter 1 and the Scientific Skills Review in Appendix F and the Study Area of MasteringBiology.)
2. Explain how the types of data the researchers chose to collect enabled them to test their prediction.
3. What conclusion do you draw from the data presented above?
4. Predict what would happen if, after 22 months, guppies from the transplanted population were returned to the source pool. Describe an experiment to test your prediction.

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

What, then, is theoretical about evolution? Keep in mind that the scientific meaning of the term *theory* is very different from its meaning in everyday use. The colloquial use of the word *theory* comes close to what scientists mean by a hypothesis. In science, a theory is more comprehensive than a hypothesis. A theory, such as the theory of evolution by natural selection, accounts for many observations and explains and integrates a great variety of phenomena. Such a unifying theory does not become widely accepted unless its predictions stand up to thorough and continual testing by experiment and additional observation (see Concept 1.3). As the rest of this unit demonstrates, this has certainly been the case with the theory of evolution by natural selection.

The skepticism of scientists as they continue to test theories prevents these ideas from becoming dogma. For example, although Darwin thought that evolution was a very slow process, we now know that this isn't always true. Populations can evolve rapidly, and new species can form in relatively short periods of time: a few thousand years or less. Furthermore, evolutionary biologists now recognize that natural selection is not the only mechanism responsible for evolution. Indeed, the study of evolution today is livelier than ever as scientists use a wide range of experimental approaches and genetic analyses to test predictions based on natural selection and other evolutionary mechanisms.

Although Darwin's theory attributes life's diversity to natural processes, the diverse products of evolution are nevertheless elegant and inspiring. As Darwin wrote in the final sentence of *The Origin of Species*, "There is grandeur in this view of life . . . [in which] endless forms most beautiful and most wonderful have been, and are being, evolved."

CONCEPT CHECK 21.3

1. Explain how the following statement is inaccurate: "Antibiotics have created drug resistance in MRSA."
2. How does evolution account for (a) the similar mammalian forelimbs with different functions shown in Figure 21.15 and (b) the similar forms of the two distantly related mammals shown in Figure 21.18?
3. **WHAT IF?** ➤ Fossils show that dinosaurs originated 200–250 million years ago. Would you expect the geographic distribution of early dinosaur fossils to be broad (on many continents) or narrow (on one or a few continents only)? Explain.

For suggested answers, see Appendix A.

21 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 21.1

The Darwinian revolution challenged traditional views of a young Earth inhabited by unchanging species *(pp. 501–503)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

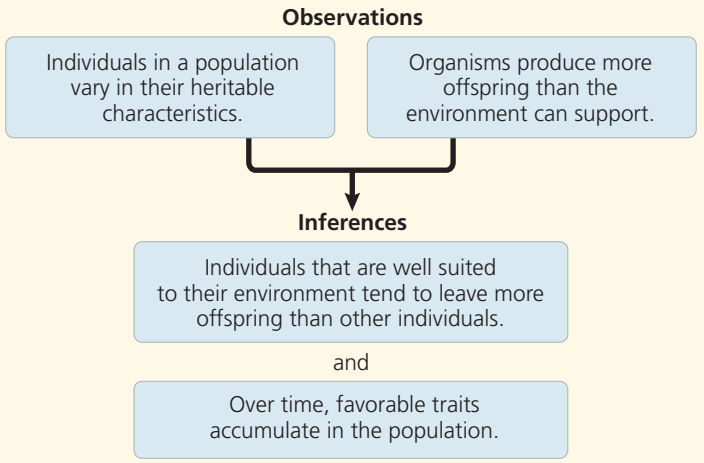
- Darwin proposed that life's diversity arose from ancestral species through natural selection, a departure from prevailing views.
- Cuvier studied **fossils** but denied that evolution occurs; he proposed that sudden catastrophic events in the past caused species to disappear from an area.
- Hutton and Lyell thought that geologic change could result from gradual mechanisms that operated in the past in the same manner as they do today.
- Lamarck hypothesized that species evolve, but the underlying mechanisms he proposed are not supported by evidence.

? *Why was the age of Earth important for Darwin's ideas about evolution?*

CONCEPT 21.2

Descent with modification by natural selection explains the adaptations of organisms and the unity and diversity of life *(pp. 503–508)*

- Darwin's experiences during the voyage of the *Beagle* gave rise to his idea that new species originate from ancestral forms through the accumulation of **adaptations**. He refined his theory for many years and finally published it in 1859 after learning that Wallace had come to the same idea.
- In *The Origin of Species*, Darwin proposed that over long periods of time, descent with modification produced the rich diversity of life through the mechanism of **natural selection**.



? *Describe how overreproduction and heritable variation relate to evolution by natural selection.*

CONCEPT 21.3

Evolution is supported by an overwhelming amount of scientific evidence *(pp. 509–516)*

- Researchers have directly observed natural selection leading to adaptive evolution in many studies, including research on soapberry bug populations and on MRSA.

- Organisms share characteristics because of common descent (**homology**) or because natural selection affects independently evolving species in similar environments in similar ways (**convergent evolution**).
- Fossils show that past organisms differed from living organisms, that many species have become extinct, and that species have evolved over long periods of time; fossils also document the evolutionary origin of new groups of organisms.
- Evolutionary theory can explain some biogeographic patterns.

Summarize the different lines of evidence supporting the hypothesis that cetaceans descended from land mammals and are closely related to even-toed ungulates.

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–5 can be found in the Study Area in MasteringBiology.

6. **EVOLUTION CONNECTION** Explain why anatomical and molecular features often fit a similar nested pattern. In addition, describe a process that can cause this not to be the case.

PRACTICE TEST
goo.gl/iAsVgL

7. **SCIENTIFIC INQUIRY • DRAW IT** Mosquitoes resistant to the pesticide DDT first appeared in India in 1959, but now are found throughout the world. (a) Graph the data in the table below. (b) Examine the graph, then hypothesize why the percentage of mosquitoes resistant to DDT rose rapidly. (c) Suggest an explanation for the global spread of DDT resistance.

Month	0	8	12
Mosquitoes Resistant* to DDT	4%	45%	77%

*Mosquitoes were considered resistant if they were not killed within 1 hour of receiving a dose of 4% DDT.

Data from C. F. Curtis et al., Selection for and against insecticide resistance and possible methods of inhibiting the evolution of resistance in mosquitoes, *Ecological Entomology* 3:273–287 (1978).

8. **WRITE ABOUT A THEME: INTERACTIONS** Write a short essay (about 100–150 words) evaluating whether changes to an organism's physical environment are likely to result in evolutionary change. Use an example to support your reasoning.

9. **SYNTHESIZE YOUR KNOWLEDGE**

This honeypot ant (genus *Myrmecocystus*) can store liquid food inside its expandable abdomen. Consider other ants you are familiar with, and explain how a honeypot ant exemplifies three key features of life: adaptation, unity, and diversity.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

This page is intentionally left blank.

22 Phylogenetic Reconstruction

Figure 22.1 What kind of organism is this?

KEY CONCEPTS

Investigating the Tree of Life

Look closely at the organism in **Figure 22.1**. Although it resembles a snake, this animal is actually a legless lizard known as the European glass lizard (*Ophisaurus apodus*). Why isn't this glass lizard considered a snake? More generally, how do biologists distinguish and categorize the millions of species on Earth?

An understanding of evolutionary relationships suggests one way to address these questions: We can decide in which category to place a species by comparing its traits with those of potential close relatives. For example, the glass lizard does not have a highly mobile jaw, a large number of vertebrae, or a short tail located behind the anus, three traits shared by all snakes. These and other characteristics suggest that despite a superficial resemblance, the glass lizard is not a snake.

Snakes and lizards are part of the continuum of life extending from the earliest organisms to the great variety of species alive today. In this unit, we will survey this diversity and describe hypotheses regarding how it evolved. As we do so, our emphasis will shift from the *process* of evolution (the evolutionary mechanisms described in Unit Four) to its *pattern* (observations of evolution's products over time).

To set the stage for surveying life's diversity, in this chapter we consider how biologists trace **phylogeny**, the evolutionary history of a species or group of species. A phylogeny of lizards and snakes, for example, indicates that both the eastern glass lizard and snakes evolved from lizards with legs—but they evolved from different

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

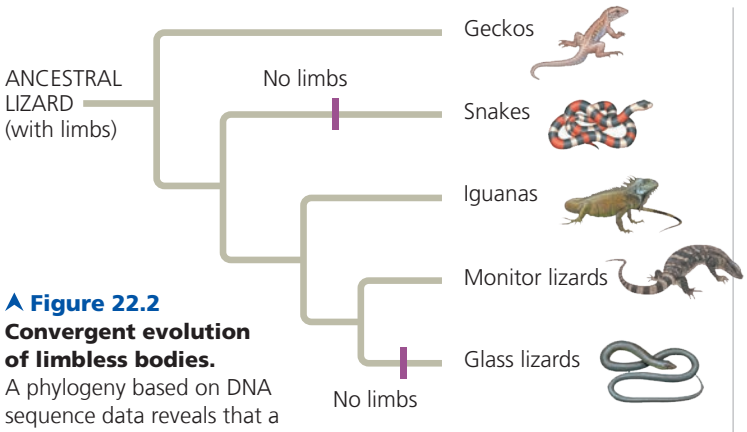


Figure 22.2
Convergent evolution of limbless bodies.
A phylogeny based on DNA sequence data reveals that a legless body form evolved independently from legged ancestors in the lineages leading to the glass lizard and to snakes.

lineages of legged lizards **(Figure 22.2)**. Thus, it appears that their legless conditions evolved independently. As we'll see, biologists reconstruct phylogenies like that in Figure 22.2 using **systematics**, a discipline focused on classifying organisms and determining their evolutionary relationships.

CONCEPT 22.1

Phylogenies show evolutionary relationships

Organisms share many characteristics because of common ancestry (see Concept 21.3). As a result, we can learn a great deal about a species if we know its evolutionary history. For example, an organism is likely to share many of its genes, metabolic pathways, and structural proteins with its close relatives. We'll consider practical applications of such information later in this section, but first we'll examine how organisms are named and classified, the scientific discipline of **taxonomy**. We'll also look at how we can interpret and use diagrams that represent evolutionary history.

Binomial Nomenclature

Common names for organisms—such as monkey, finch, and lilac—convey meaning in casual usage, but they can also cause confusion. Each of these names, for example, refers to more than one species. Moreover, some common names do not accurately reflect the kind of organism they signify. Consider these three "fishes": jellyfish (a cnidarian), crayfish (a small, lobsterlike crustacean), and silverfish (an insect). And, of course, a given organism has different names in different languages.

To avoid ambiguity when communicating about their research, biologists refer to organisms by Latin scientific names. The two-part format of the scientific name, commonly called a **binomial**, was instituted in the 18th century by Carolus Linnaeus (see Concept 21.1). The first part of a binomial is the name of the **genus** (plural, *genera*) to which the species belongs. The second part, called the specific epithet, is unique for each species within the genus. An example of a binomial is *Panthera pardus*, the scientific name for the leopard. Notice that the first letter of the genus is capitalized and the entire binomial is italicized. (Newly created scientific names are also "latinized": You can name an insect you discover after a friend, but you must add a Latin ending.) Many of the more than 11,000 binomials assigned by Linnaeus are still used today, including the optimistic name he gave our own species—*Homo sapiens*, meaning "wise man."

Hierarchical Classification

In addition to naming species, Linnaeus also grouped them into a hierarchy of increasingly inclusive categories. The first grouping is built into the binomial: Species that appear to be closely related are grouped into the same genus. For example, the leopard (*Panthera pardus*) belongs to a genus that also includes the African lion (*Panthera leo*), the tiger (*Panthera tigris*), and the jaguar (*Panthera onca*). Beyond genera, taxonomists employ progressively more comprehensive categories of classification. The taxonomic system named after Linnaeus, the Linnaean system, places related genera in the same **family**, families into **orders**, orders into **classes**, classes into **phyla** (singular, *phylum*), phyla into **kingdoms**, and, more recently, kingdoms into **domains (Figure 22.3)**. The resulting biological classification of a particular organism is somewhat like a postal address identifying a person in a particular apartment, in a building with many apartments, on a street with many apartment buildings, in a city with many streets, and so on.

The named taxonomic unit at any level of the hierarchy is called a **taxon** (plural, *taxa*). In the leopard example, *Panthera* is a taxon at the genus level, and Mammalia is a taxon at the class level that includes all the many orders of mammals. Note that in the Linnaean system, taxa broader than the genus are not italicized, though they are capitalized.

Classifying species is a way to structure our human view of the world. We lump together various species of trees to which we give the common name of pines and distinguish them from other trees that we call firs. Taxonomists have decided that pines and firs are different enough to be placed in separate genera, yet similar enough to be grouped into the same family, Pinaceae. As with pines and firs, higher levels of classification are usually defined by particular characters chosen by taxonomists. However, characters that are useful for classifying one group of organisms may not be appropriate for other organisms. For this reason, the larger categories often are not comparable between lineages; that is, an order of snails does not exhibit the same degree of morphological or genetic diversity as an order of mammals. As we'll see, the placement of species into orders, classes, and so on also does not necessarily reflect evolutionary history.

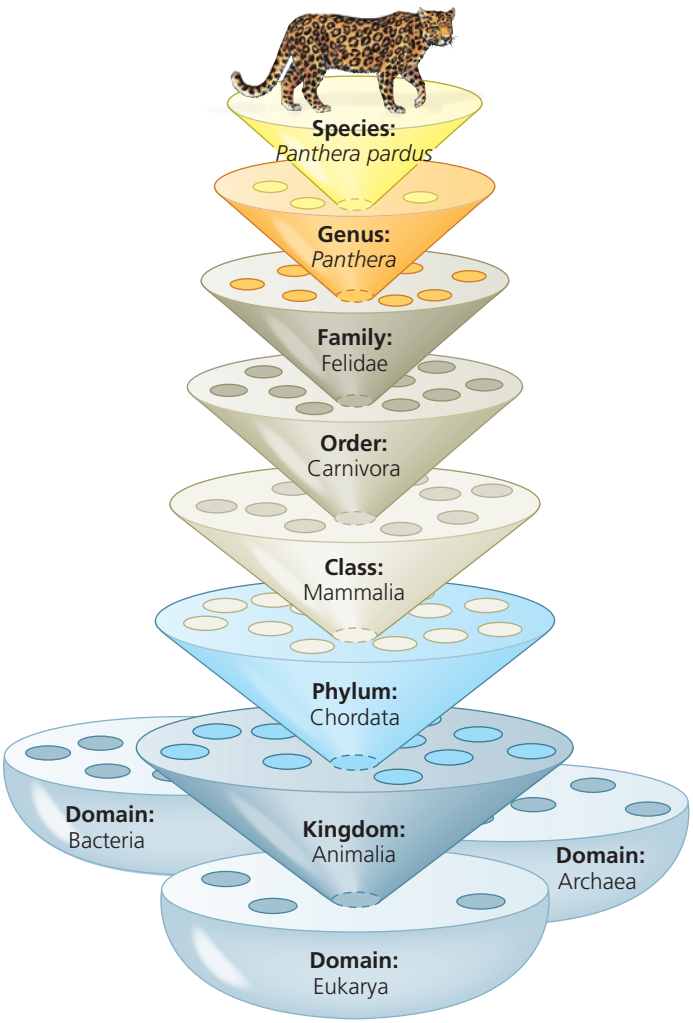


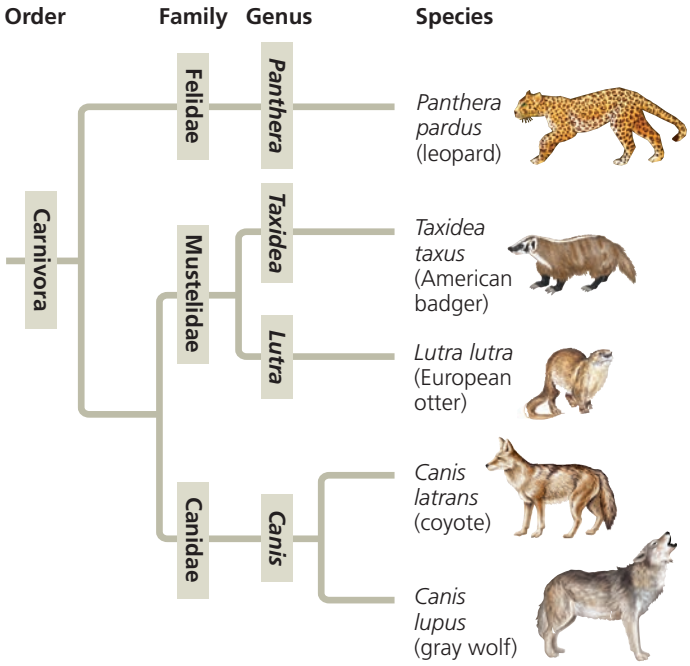
Figure 22.3 Linnaean classification. At each level, or "rank," species are placed in groups within more inclusive groups.

Linking Classification and Phylogeny

The evolutionary history of a group of organisms can be represented in a branching diagram called a **phylogenetic tree**. As in **Figure 22.4**, the branching pattern often matches how taxonomists have classified groups of organisms nested within more inclusive groups. Sometimes, however, taxonomists have placed a species within a genus (or other group) to which it is *not* most closely related. One reason for such a mistake might be that over the course of evolution, a species has lost a key feature shared by its close relatives. If DNA or other new evidence indicates that an organism has been misclassified, the organism may be reclassified to accurately reflect its evolutionary history. Another issue is that while the Linnaean system may distinguish groups, such as amphibians, mammals, reptiles, and other classes of vertebrates, it tells us nothing about these groups' evolutionary relationships to one another.

Such difficulties in aligning Linnaean classification with phylogeny have led some systematists to propose that

Figure 22.4 The connection between classification and phylogeny. Hierarchical classification can reflect the branching patterns of phylogenetic trees. This tree shows evolutionary relationships between some of the taxa within order Carnivora, itself a branch of class Mammalia.



classification be based entirely on evolutionary relationships. In such systems, names are only assigned to groups that include a common ancestor and all of its descendants. As a consequence of this approach, some commonly recognized groups would become part of other groups previously at the same level of the Linnaean system. For example, because birds evolved from a group of reptiles, Aves (the Linnaean class to which birds are assigned) would be considered a subgroup of Reptilia (also a class in the Linnaean system).

What We Can and Cannot Learn from Phylogenetic Trees

Regardless of how groups are named, a phylogenetic tree represents a hypothesis about evolutionary relationships **(Figure 22.5)**. These relationships often are depicted as a series of dichotomies, or two-way branch points. Each **branch point** represents the common ancestor of the two evolutionary lineages diverging from it.

In Figure 22.5, each tree has a branch point that represents the common ancestor of the lineages leading to chimpanzees and humans. Chimps and humans are considered **sister taxa**, groups of organisms that share an immediate common ancestor that is not shared by any other group. The members of a sister group are each other's closest relatives, making sister groups a useful way to describe the evolutionary relationships shown in a tree. For example, in Figure 22.5, the evolutionary lineage leading to lizards shares an immediate common ancestor with the lineage

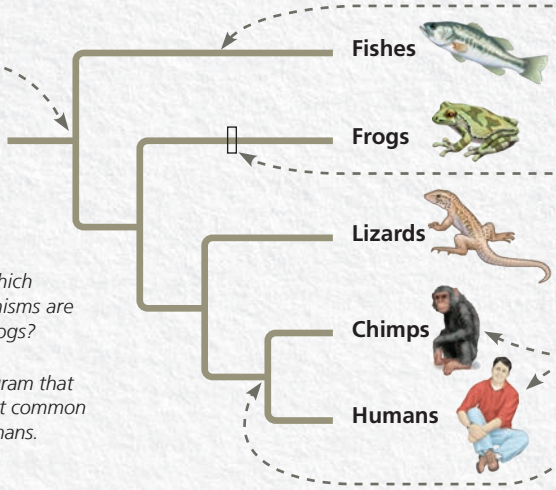
Figure 22.5 Visualizing Phylogenetic Relationships

A phylogenetic tree visually represents a hypothesis of how a group of organisms are related. This figure explores how the way a tree is drawn conveys information.

Instructors: Additional questions related to this Visualizing Figure can be assigned in MasteringBiology.

Parts of a Tree

This tree shows how the five groups of organisms at the tips of the branches, called *taxa*, are related. Each **branch point** represents the common ancestor of the evolutionary lineages diverging from it.



This branch point represents the common ancestor of all the animal groups shown in this tree.

Each horizontal branch represents an **evolutionary lineage.** The length of the branch is arbitrary unless the diagram specifies that branch lengths represent information such as time or amount of genetic change (see Figure 22.13).

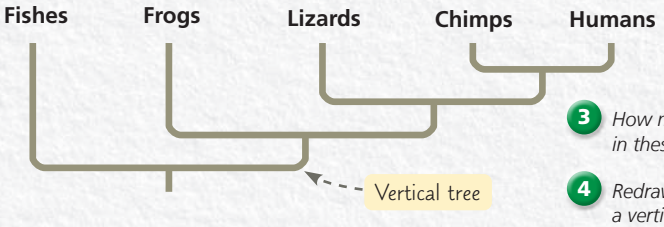
Each position along a branch represents an ancestor in the lineage leading to the taxon named at the tip.

Sister taxa are groups of organisms that share a common ancestor that is not shared by any other group. Chimps and humans are an example of sister taxa in this tree.

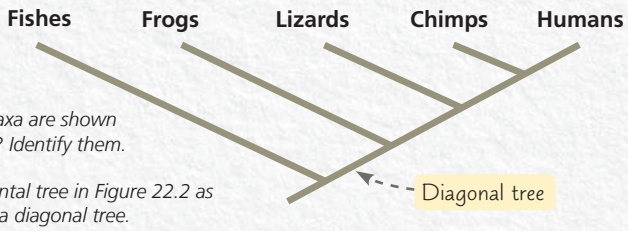
1. *According to this tree, which group or groups of organisms are most closely related to frogs?*
2. *Label the part of the diagram that represents the most recent common ancestor of frogs and humans.*

Alternative Forms of Tree Diagrams

These diagrams are referred to as "trees" because they use the visual analogy of branches to represent evolutionary lineages diverging over time. In this text, trees are usually drawn horizontally, as shown above, but the same tree can be drawn vertically or diagonally without changing the relationships it conveys.



Vertical tree

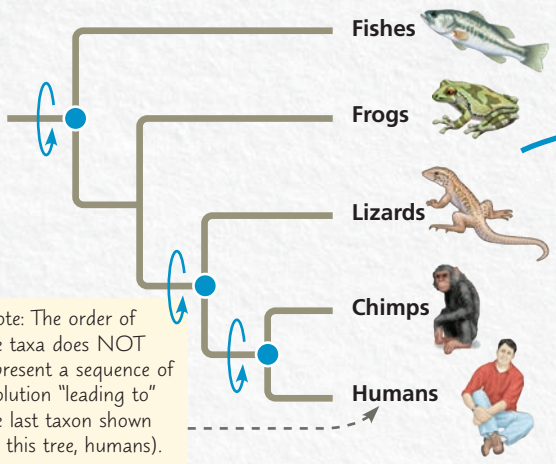


Diagonal tree

3. *How many sister taxa are shown in these two trees? Identify them.*
4. *Redraw the horizontal tree in Figure 22.2 as a vertical tree and a diagonal tree.*

Rotating Around Branch Points

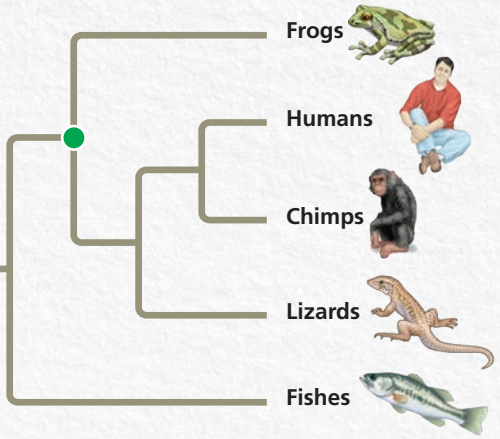
Rotating the branches of a tree around a branch point does not change what they convey about evolutionary relationships. As a result, the order in which taxa appear at the branch tips is not significant. What matters is the branching pattern, which signifies the order in which the lineages have diverged from common ancestors.



Note: The order of the taxa does NOT represent a sequence of evolution "leading to" the last taxon shown (in this tree, humans).

Rotating the branches in the tree at left around the three blue branch points yields the tree on the right.

5. *Redraw the tree on the right, rotating around the green branch point. Identify the two closest relatives of humans as shown in each of the three trees. Explain your answer.*



leading to chimpanzees and humans. Thus, we can describe this portion of the tree by saying that of the groups shown here, lizards are the sister taxon to a group consisting of chimpanzees and humans.

As also shown in Figure 22.5, the branches of a tree can be rotated around branch points without changing the relationships shown in the tree. That is, the order in which the taxa appear at the right side of the tree does not represent a *sequence* of evolution—in this case, it does not imply a sequence leading from fishes to humans.

This tree, like all of the phylogenetic trees in this book, is **rooted**, which means that a branch point within the tree (often drawn farthest to the left) represents the most recent common ancestor of all taxa in the tree. A lineage that diverges from all other members of its group early in the history of the group is called a **basal taxon**. Hence, like the fishes in Figure 22.5, a basal taxon lies on a branch that diverges near the common ancestor of the group.

What other key points do we need to keep in mind when interpreting phylogenetic trees? First, they are intended to show patterns of descent, not phenotypic similarity. Although closely related organisms often resemble one another due to their common ancestry, they may not if their lineages have evolved at different rates or faced very different environmental conditions. For example, even though crocodiles are more closely related to birds than to lizards (see Figure 21.17), they look more like lizards because morphology has changed dramatically in the bird lineage.

Second, we cannot necessarily infer the ages of the taxa or branch points shown in a tree. For example, the tree in Figure 22.5 does not indicate that chimpanzees evolved before humans. Rather, the tree shows only that chimpanzees and humans share a recent common ancestor, but we cannot tell when that ancestor lived or when the first chimpanzees or humans arose. Generally, unless given specific information about what the branch lengths in a tree mean—for example, that they are proportional to time—we should interpret the diagram solely in terms of patterns of descent. No assumptions should be made about when particular species evolved or how much change occurred in each lineage.

Third, we should not assume that a taxon on a phylogenetic tree evolved from the taxon next to it. Figure 22.5 does not indicate that humans evolved from chimpanzees or vice versa. We can infer only that the lineage leading to humans and the lineage leading to chimpanzees both evolved from a recent common ancestor. That ancestor, which is now extinct, was neither a human nor a chimpanzee.

Applying Phylogenies

Understanding phylogeny can have practical applications. Consider maize (corn), which originated in the Americas and is now an important food crop worldwide. From a phylogeny of maize based on DNA data, researchers have been able to identify two species of wild grasses that may be maize's closest living relatives. These two close relatives may be useful as "reservoirs" of beneficial alleles that can be transferred to cultivated maize by cross-breeding or genetic engineering.

A different use of phylogenetic trees is to infer species identities by analyzing the relatedness of DNA sequences from different organisms. Researchers have used this approach to investigate whether "whale meat" had been harvested illegally from whale species protected under international law rather than from species that can be harvested legally **(Figure 22.6)**.

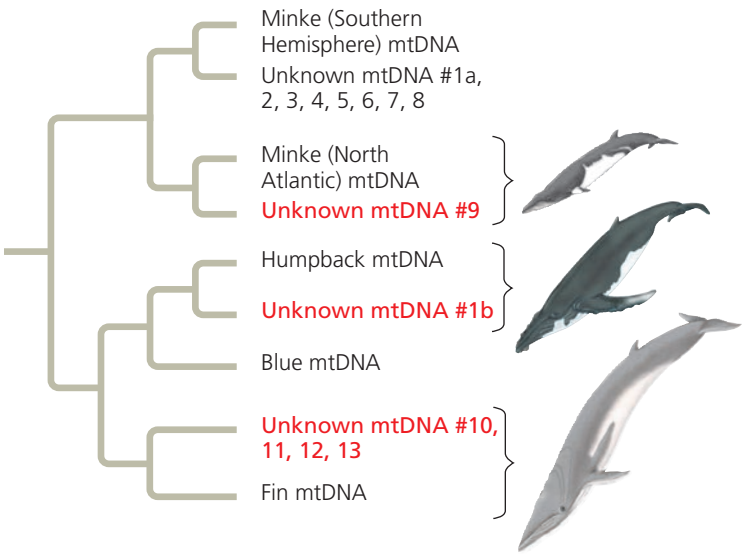
How do researchers construct phylogenetic trees like those we've considered here? In the next section, we'll begin to answer that question by examining the data used to determine phylogenies.

Figure 22.6

Inquiry What is the species identity of food being sold as whale meat?

Experiment C. S. Baker and S. R. Palumbi purchased 13 samples of "whale meat" from Japanese fish markets. They sequenced part of the mitochondrial DNA (mtDNA) from each sample and compared their results with the comparable mtDNA sequence from known whale species. To infer the species identity of each sample, the team constructed a *gene tree*, a phylogenetic tree that shows patterns of relatedness among DNA sequences rather than among taxa.

Results Of the species in the resulting gene tree, only Minke whales caught in the Southern Hemisphere can be sold legally in Japan.



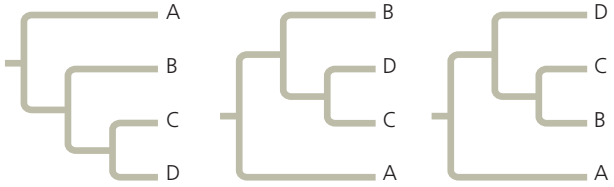
Conclusion This analysis indicated that mtDNA sequences of six of the unknown samples (in red) were most closely related to mtDNA sequences of whales that are not legal to harvest.

Data from C. S. Baker and S. R. Palumbi, Which whales are hunted? A molecular genetic approach to monitoring whaling, *Science* 265:1538–1539 (1994). Reprinted with permission from AAAS.

WHAT IF? ➤ *What different results would have indicated that all of the whale meat had been harvested legally?*

CONCEPT CHECK 22.1

1. **VISUAL SKILLS ➤** Which levels of the classification in Figure 22.3 do humans share with leopards?
2. **VISUAL SKILLS ➤** Which of the trees shown here depicts an evolutionary history different from the other two? Explain.



(a) (b) (c)

3. **DRAW IT ➤** The bear family (Ursidae) is more closely related to the badger/otter family (Mustelidae) than to the dog family (Canidae). Use this information to redraw Figure 22.4.

For suggested answers, see Appendix A.

CONCEPT 22.2

Phylogenies are inferred from morphological and molecular data

To infer phylogeny, systematists must gather as much information as possible about the morphology, genes, and biochemistry of the relevant organisms. It is important to focus on features that result from common ancestry because only those features reflect evolutionary relationships.

Morphological and Molecular Homologies

Recall that phenotypic and genetic similarities due to shared ancestry are called **homologies**. For example, the similarity in the number and arrangement of bones in the forelimbs of mammals is due to their descent from a common ancestor with the same bone structure; this is an example of a morphological homology (see Figure 21.15). In the same way, genes or other DNA sequences are homologous if they are descended from sequences carried by a common ancestor.

In general, organisms that share very similar morphologies or similar DNA sequences are likely to be more closely related than organisms with vastly different structures or sequences. In some cases, however, the morphological divergence between related species can be great and their genetic divergence small (or vice versa). Consider the Hawaiian silversword plants: some of these species are tall, twiggy trees, while others are dense, ground-hugging shrubs (see Figure 25.22). But despite these striking phenotypic differences, the silverswords' genes are very similar. Based on these small molecular divergences, scientists estimate that the silversword group began to diverge 5 million years ago. We'll discuss how scientists use molecular data to estimate such divergence times later in this chapter.

Sorting Homology from Analogy

A potential source of confusion in constructing a phylogeny is similarity between organisms that is due to convergent evolution—called **analogy**—rather than to shared ancestry (homology). Convergent evolution occurs when similar environmental pressures and natural selection produce similar (analogous) adaptations in organisms from different evolutionary lineages.

Figure 22.7 Convergent evolution in burrowers. A long body, large front paws, small eyes, and a pad of thick skin that protects the nose all evolved independently in these species.

Australian "mole"

African golden mole

For example, the two mole-like animals shown in **Figure 22.7** look very similar. However, their internal anatomy, physiology, and reproductive systems are very dissimilar. Indeed, genetic and fossil evidence indicate that the common ancestor of these animals lived 140 million years ago. This common ancestor and most of its descendants were not mole-like. It appears that analogous characteristics evolved independently in these two lineages as they became adapted to similar lifestyles—hence, the similar features of these animals should not be considered when reconstructing their phylogeny.

Another clue to distinguishing between homology and analogy is the complexity of the characters being compared. The more elements that are similar in two complex structures, the more likely it is that the structures evolved from a common ancestor. For instance, the skulls of an adult human and an adult chimpanzee both consist of many bones fused together. The compositions of the skulls match almost perfectly, bone for bone. It is highly improbable that such complex structures, matching in so many details, have separate origins. More likely, the genes involved in the development of both skulls were inherited from a common ancestor.

The same argument applies to comparisons at the gene level. Genes are sequences of thousands of nucleotides, each of which represents an inherited character in the form of one of the four DNA bases: A (adenine), G (guanine), C (cytosine), or T (thymine). If genes in two organisms share many portions of their nucleotide sequences, it is likely that the genes are homologous.

Evaluating Molecular Homologies

Comparing DNA molecules often poses technical challenges for researchers. The first step after sequencing the molecules is to align comparable sequences from the species being studied. If the species are very closely related, the sequences probably differ at only one or a few sites. In contrast, comparable nucleic acid sequences in distantly related species

usually have different bases at many sites and may have different lengths. This is because insertions and deletions accumulate over long periods of time.

Suppose, for example, that certain noncoding DNA sequences near a particular gene are very similar in two species, except that the first base of the sequence has been deleted in one of the species. The effect is that the remaining sequence shifts back one notch. A comparison of the two sequences that does not take this deletion into account would overlook what in fact is a very good match. To address such problems, researchers have developed computer programs that estimate the best way to align comparable DNA segments of differing lengths **(Figure 22.8)**.

Such molecular comparisons reveal that many base substitutions and other differences have accumulated in the comparable genes of an Australian "mole" and a golden mole. The many differences indicate that their lineages have diverged greatly since their common ancestor; thus, we say that the living species are not closely related. In contrast, the high degree of gene sequence similarity among the silversword plants indicates that they are all very closely related, in spite of their considerable morphological differences.

Just as with morphological characters, it is necessary to distinguish homology from analogy in evaluating molecular similarities for evolutionary studies. Two sequences that resemble each other at many points along their length most likely are homologous (see Figure 22.8). But in organisms that do not appear to be closely related, the bases that their otherwise very different sequences happen to share may simply be coincidental matches, called molecular homoplasies. For example, if the two DNA sequences in **Figure 22.9** were from distantly related organisms, the fact that they share 23% of their bases would be coincidental. Statistical tools have been developed to determine whether DNA sequences that share more than 25% of their bases do so because they are homologous.

Figure 22.8 Aligning segments of DNA. Systematists search for similar sequences along DNA segments from two species (only one DNA strand is shown for each species). In this example, 11 of the original 12 bases have not changed since the species diverged. Hence, those portions of the sequences still align once the length is adjusted.

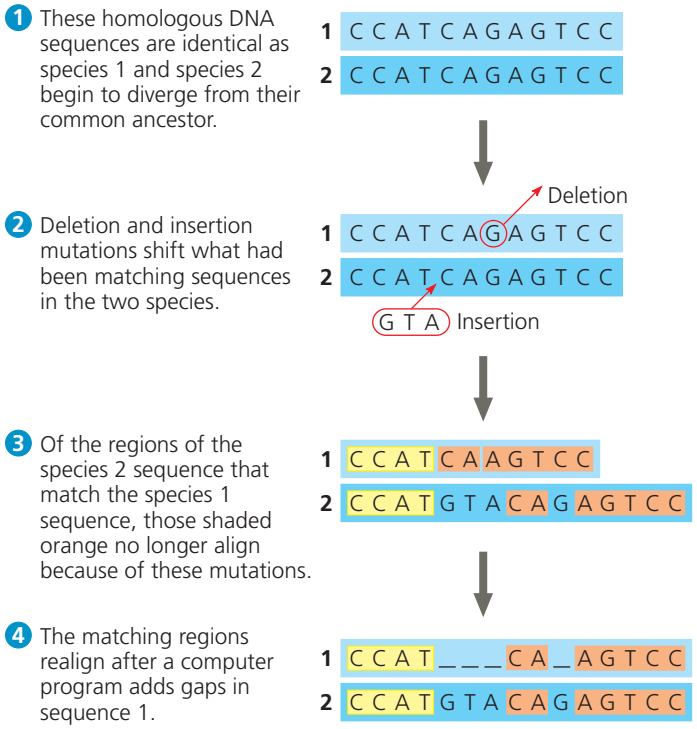
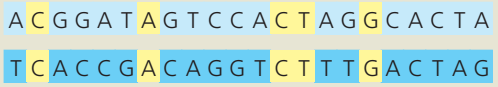


Figure 22.9 A molecular homoplasy.



CONCEPT CHECK 22.2

1. Decide whether each of the following pairs of structures more likely represents analogy or homology, and explain your reasoning: (a) a porcupine's quills and a cactus's spines; (b) a cat's paw and a human's hand; (c) an owl's wing and a hornet's wing.
2. **WHAT IF?** ➤ Suppose that two species, A and B, have similar appearances but very divergent gene sequences, while species B and C have very different appearances but similar gene sequences. Which pair of species is more likely to be closely related: A and B or B and C? Explain.

For suggested answers, see Appendix A.

CONCEPT 22.3

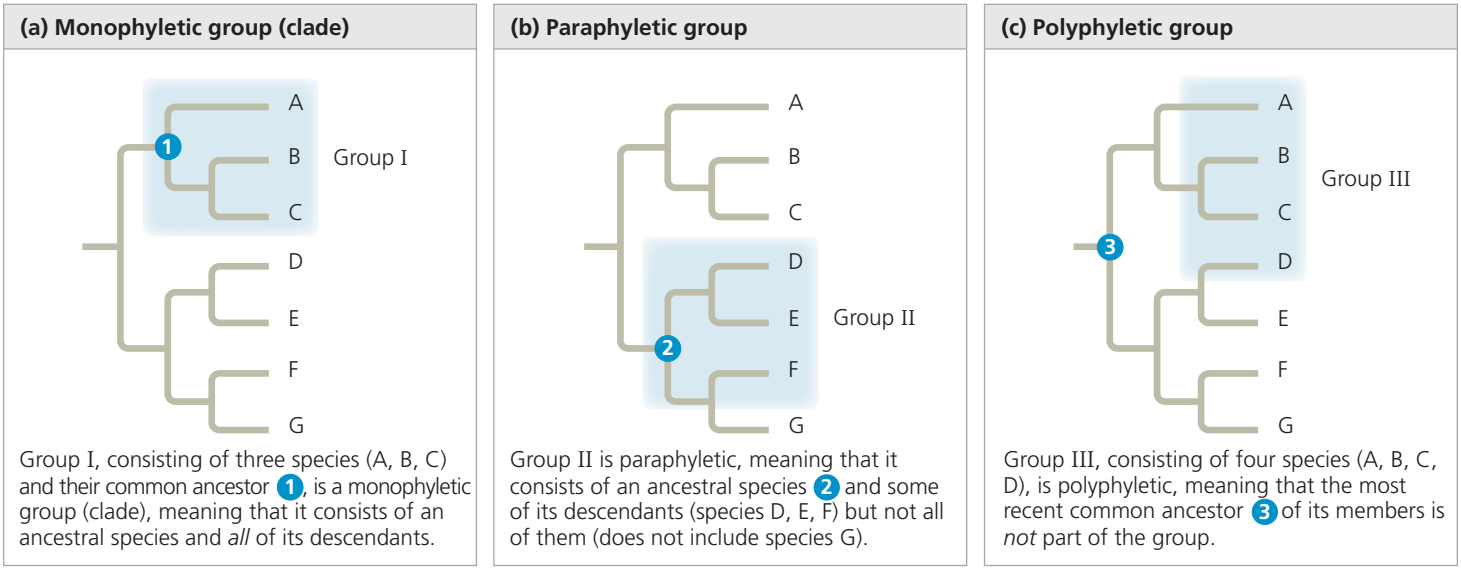
Shared characters are used to construct phylogenetic trees

As we've discussed, a key step in reconstructing phylogenies is to distinguish homologous features from analogous ones (since only homology reflects evolutionary history). Next, we'll describe cladistics, a widely used set of methods for inferring phylogeny from homologous characters.

Cladistics

In the approach to systematics called **cladistics**, common ancestry is the primary criterion used to classify organisms. Using this methodology, biologists attempt to place species into groups called **clades**, each of which includes an ancestral species and all of its descendants. Clades, like taxonomic categories of the Linnaean system, are nested within larger clades. In Figure 22.4, for example, the cat group (Felidae) represents a clade within a larger clade (Carnivora) that also includes the dog group (Canidae).

Figure 22.10 Monophyletic, paraphyletic, and polyphyletic groups.



However, a taxon is equivalent to a clade only if it is **monophyletic** (from the Greek, meaning "single tribe"), signifying that it consists of an ancestral species and all of its descendants **(Figure 22.10a)**. Contrast this with a **paraphyletic** ("beside the tribe") group, which consists of an ancestral species and some, but not all, of its descendants **(Figure 22.10b)**, or a **polyphyletic** ("many tribes") group, which includes distantly related species but does not include their most recent common ancestor **(Figure 22.10c)**.

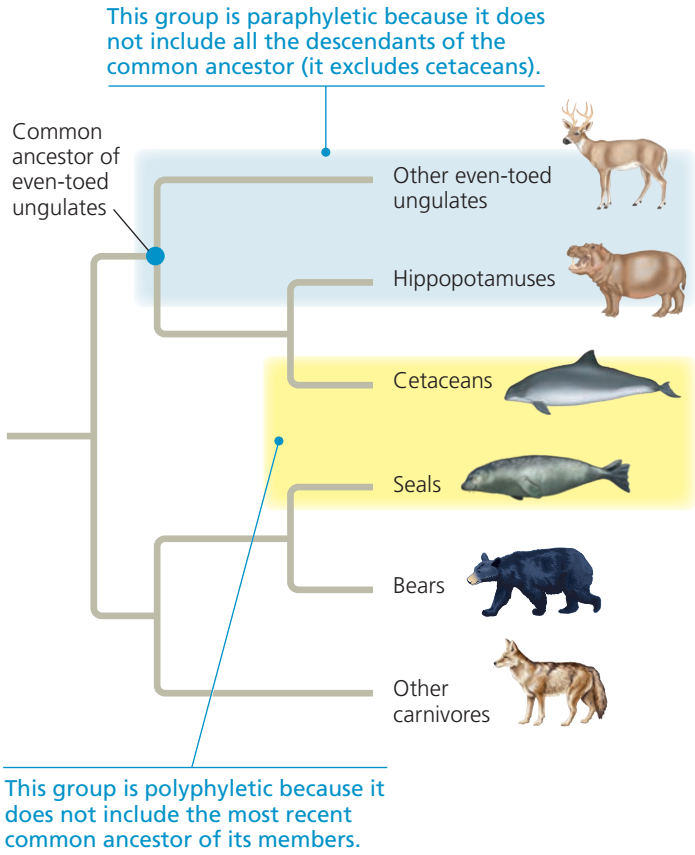
Note that in a paraphyletic group, the most recent common ancestor of all members of the group *is* part of the group, whereas in a polyphyletic group the most recent common ancestor *is not* part of the group. For example, a group consisting of even-toed ungulates (hippopotamuses, deer, and their relatives) and their common ancestor is paraphyletic because it does not include cetaceans (whales, dolphins, and porpoises), which descended from that ancestor **(Figure 22.11)**. In contrast, a group consisting of seals and cetaceans (based on their similar body forms) is polyphyletic because it does not include the common ancestor of seals and cetaceans. Biologists avoid defining such polyphyletic groups; if new evidence indicates that an existing group is polyphyletic, its members are reclassified.

Shared Ancestral and Shared Derived Characters

As a result of descent with modification, organisms have characters they share with their ancestors, and they also have characters that differ from those of their ancestors. For example, all mammals have backbones, but a backbone does not distinguish mammals from other vertebrates because *all* vertebrates have backbones. The backbone predates the branching of mammals from other vertebrates. Thus for mammals, the backbone is a **shared ancestral character**, a character that originated in an ancestor of the taxon.

In contrast, hair is a character shared by all mammals but *not* found in their ancestors. Thus, in mammals, hair is considered a **shared derived character**, an evolutionary novelty unique to a clade.

Figure 22.11 Examples of a paraphyletic and a polyphyletic group.



DRAW IT ➤ *Circle the branch point that represents the most recent common ancestor of cetaceans and seals. Explain why that ancestor would not be part of a cetacean–seal group defined by their similar body forms.*

Note that a shared derived character can refer to the loss of a feature, such as the loss of limbs in snakes or whales. In addition, it is a relative matter whether a character is considered ancestral or derived. A backbone can also qualify as a shared derived character, but only at a deeper branch point that distinguishes all vertebrates from other animals.

Inferring Phylogenies Using Derived Characters

Shared derived characters are unique to particular clades. Because all features of organisms arose at some point in the history of life, it should be possible to determine the clade in which each shared derived character first appeared and to use that information to infer evolutionary relationships.

To give an example of this approach, consider the set of characters shown in **Figure 22.12a** for each of five vertebrates—a leopard, turtle, frog, bass, and lamprey (a jawless aquatic vertebrate). As a basis of comparison, we need to select an outgroup. An **outgroup** is a species or group of species from an evolutionary lineage that is closely related to but not part of the group of species that we are studying (the **ingroup**). A suitable outgroup can be determined based on evidence from morphology, paleontology, embryonic development, and gene sequences. An appropriate outgroup for our example is the lancelet, a small animal that lives in mudflats and (like vertebrates) is a member of the more inclusive group called the chordates. Unlike the vertebrates, however, the lancelet does not have a backbone.

In our analysis, a character found in both the outgroup and the ingroup is assumed to be ancestral. We'll also assume that each derived character in Figure 22.12a arose only once in the ingroup. Thus, for a character that only occurs in a subset of the ingroup, we'll assume that the character arose in the lineage leading to those members of the ingroup.

By comparing members of the ingroup with each other and with the outgroup, we can determine which characters were derived at the various branch points of vertebrate evolution. In our example, *all* of the vertebrates in the ingroup have backbones: This character was present in the ancestral vertebrate, but not in the outgroup. Now note that hinged jaws are absent in the outgroup and in lampreys, but present in all other members of the ingroup. This indicates that hinged jaws arose in a lineage leading to all members of the ingroup *except* lampreys. Hence, we can conclude that lampreys are the sister taxon to the other vertebrates in the ingroup. Proceeding in this way, we can translate the data in our table of characters into a phylogenetic tree that places all the ingroup taxa into a hierarchy based on their shared derived characters **(Figure 22.12b)**.

Phylogenetic Trees with Proportional Branch Lengths

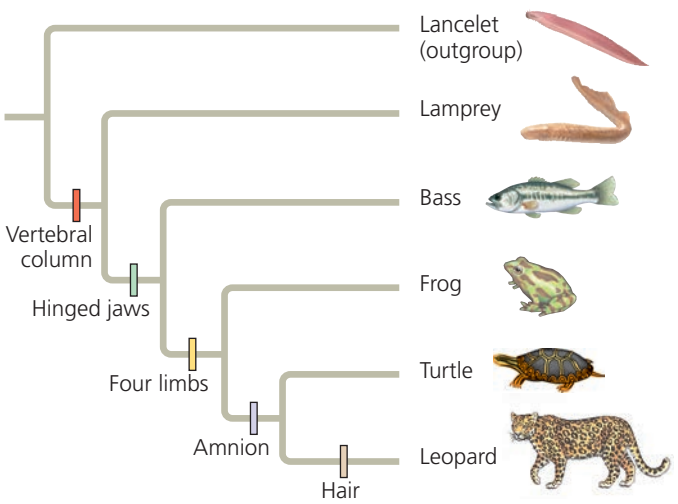
In the phylogenetic trees we have presented so far, the lengths of the tree's branches do not indicate the degree of evolutionary change in each lineage. Furthermore, the chronology represented by the branching pattern of the tree is relative (earlier versus later) rather than absolute (how many millions of years ago). But in some tree diagrams, branch lengths are proportional to amount of evolutionary change or to the times at which particular events occurred.

Figure 22.12 Using derived characters to infer phylogeny. The derived characters used here include the amnion, a membrane that encloses the embryo inside a fluid-filled sac (see Figure 34.26). Note that a different set of characters could lead us to infer a different phylogenetic tree.

TAXA

CHARACTERS	Lancelet (outgroup)	Lamprey	Bass	Frog	Turtle	Leopard
Vertebral column (backbone)	0	1	1	1	1	1
Hinged jaws	0	0	1	1	1	1
Four limbs	0	0	0	1	1	1
Amnion	0	0	0	0	1	1
Hair	0	0	0	0	0	1

(a) Character table. A 0 indicates that a character is absent; a 1 indicates that a character is present.



(b) Phylogenetic tree. The character table in (a) leads us to infer the phylogenetic relationships shown in this tree.

DRAW IT ➤ *In (b), circle the most inclusive clade for which a hinged jaw is a shared* ancestral *character.*

MB **Figure Walkthrough**

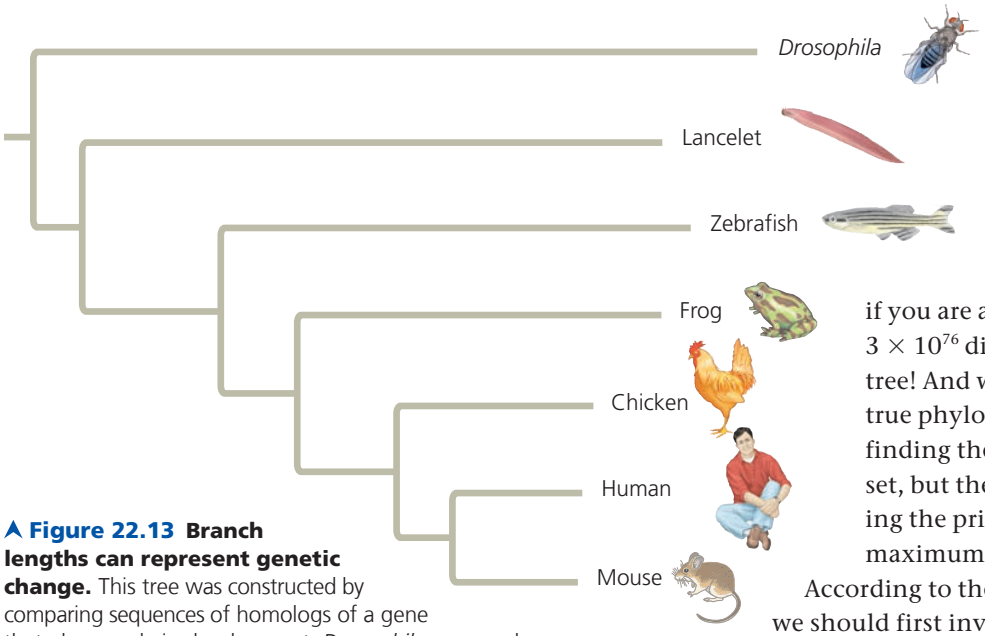


Figure 22.13 Branch lengths can represent genetic change. This tree was constructed by comparing sequences of homologs of a gene that plays a role in development; *Drosophila* was used as an outgroup. The branch lengths are proportional to the amount of genetic change in each lineage; varying branch lengths indicate that the gene has evolved at different rates in different lineages.

INTERPRET THE DATA *In which vertebrate lineage has the studied gene evolved most rapidly? Explain.*

In **Figure 22.13**, for example, the branch length of the phylogenetic tree reflects the number of changes that have taken place in a particular DNA sequence in that lineage. Note that the total length of the horizontal lines from the base of the tree to the mouse is less than that of the line leading to the outgroup species, the fruit fly *Drosophila*. This implies that in the time since the mouse and fly lineages diverged from their common ancestor, more genetic changes have occurred in the *Drosophila* lineage than in the mouse lineage.

Even though the branches of a phylogenetic tree may have different lengths, among organisms alive today, all the different lineages that descend from a common ancestor have survived for the same number of years. To take an extreme example, humans and bacteria had a common ancestor that lived over 3 billion years ago. Fossils and genetic evidence indicate that this ancestor was a single-celled prokaryote. Even though bacteria have apparently changed little in their morphology since that common ancestor, there have nonetheless been 3 billion years of evolution in the bacterial lineage, just as there have been 3 billion years of evolution in the lineage that ultimately gave rise to humans.

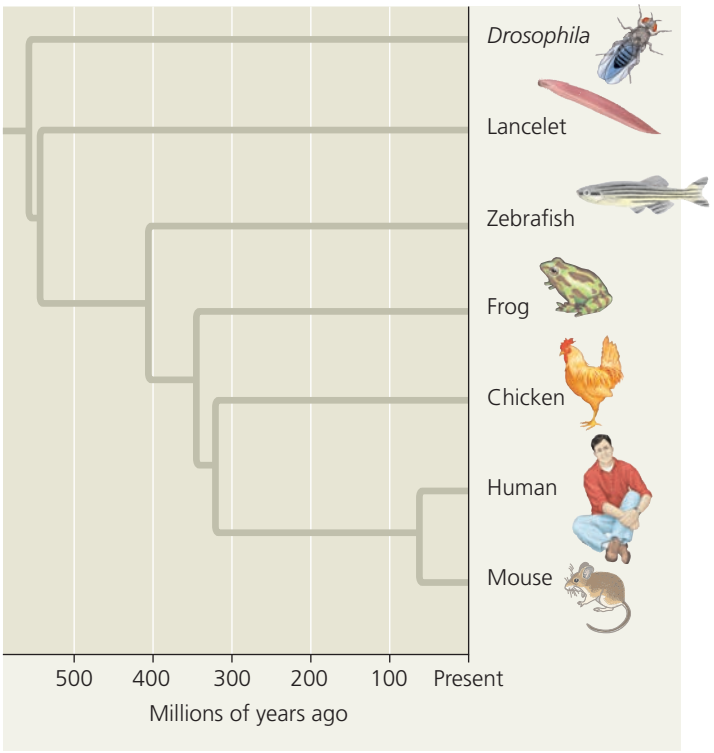
These equal spans of chronological time can be represented in a phylogenetic tree whose branch lengths are proportional to time **(Figure 22.14)**. Such a tree draws on fossil data to place branch points in the context of geologic time. Additionally, it is possible to combine these two types of trees by labeling branch points with information about rates of genetic change or dates of divergence.

Maximum Parsimony and Maximum Likelihood

As the database of DNA sequences that enables us to study more species grows, the difficulty of building the phylogenetic tree that best describes their evolutionary history also grows. What if you are analyzing data for 50 species? There are 3×10^{76} different ways to arrange 50 species into a tree! And which tree in this huge forest reflects the true phylogeny? Systematists can never be sure of finding the most accurate tree in such a large data set, but they can narrow the possibilities by applying the principles of maximum parsimony and maximum likelihood.

According to the principle of **maximum parsimony**, we should first investigate the simplest explanation that is consistent with the facts. (The parsimony principle is also called "Occam's razor" after William of Occam, a 14th-century English philosopher who advocated this minimalist problem-solving approach of "shaving away" unnecessary complications.) In the case of trees based on morphology, the most parsimonious tree requires the fewest evolutionary events, as measured by the origin of shared

Figure 22.14 Branch lengths can indicate time. This tree is based on the same DNA data as that in Figure 22.13, but here the branch points are dated based on fossil evidence. Thus, the branch lengths are proportional to time. Each lineage has the same total length from the base of the tree to the branch tip, indicating that all the lineages have diverged from the common ancestor for equal amounts of time.



derived morphological characters. For phylogenies based on DNA, the most parsimonious tree requires the fewest base changes.

A **maximum likelihood** approach identifies the tree most likely to have produced a given set of DNA data, based on certain probability rules about how DNA sequences change over time. For example, the underlying probability rules could be based on the assumption that all nucleotide substitutions are equally likely. However, if evidence suggests that this assumption is not correct, more complex rules could be devised to account for different rates of change among different nucleotides or at different positions in a gene.

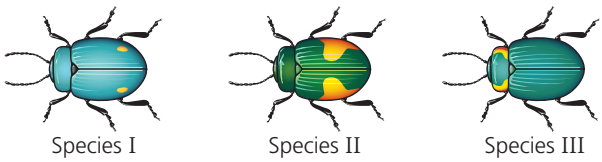
Scientists have developed many computer programs to search for trees that are parsimonious and likely. When a large amount of accurate data is available, the methods used in these programs usually yield similar trees. As an example of one method, **Figure 22.15** walks you through the process of identifying the most parsimonious molecular tree for

Figure 22.15

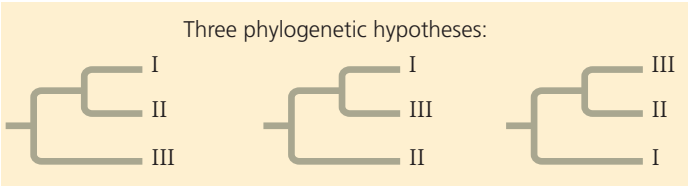
Research Method Applying Parsimony to a Problem in Molecular Systematics

Application In considering possible phylogenies for a group of species, systematists compare molecular data for the species. An efficient way to begin is by identifying the most parsimonious hypothesis—the one that requires the fewest evolutionary events (molecular changes) to have occurred.

Technique Follow the numbered steps as we apply the principle of parsimony to a hypothetical phylogenetic problem involving three closely related beetle species.



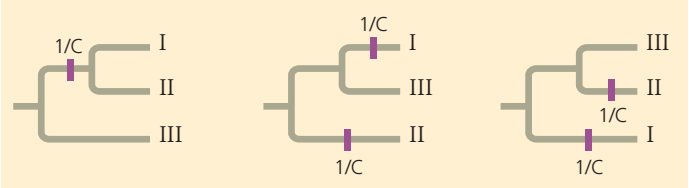
1 First, draw the three possible trees for the species. (Although only 3 trees are possible when ordering 3 species, the number of possible trees increases rapidly with the number of species: There are 15 trees for 4 species and 34,459,425 trees for 10 species.)



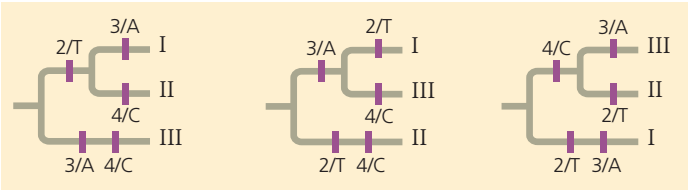
2 Tabulate the molecular data for the species. In this simplified example, the data represent a DNA sequence consisting of just four nucleotide bases. Data from several outgroup species (not shown) were used to infer the ancestral DNA sequence.

	Site 1	2	3	4
Species I	C	T	A	T
Species II	C	T	T	C
Species III	A	G	A	C
Ancestral sequence	A	G	T	T

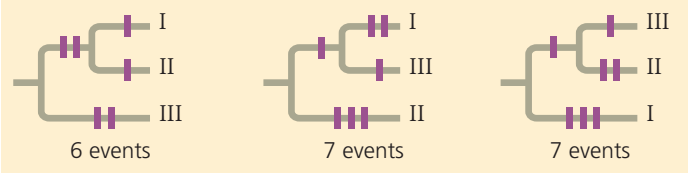
3 Now focus on site 1 in the DNA sequence. In the tree on the left, a single base-change event, represented by the purple hatch mark on the branch leading to species I and II (and labeled 1/C, indicating a change at site 1 to nucleotide C), is sufficient to account for the site 1 data. In the other two trees, two base-change events are necessary.



4 Continuing the comparison of bases at sites 2, 3, and 4 reveals that each of the three trees requires a total of five additional base-change events (purple hatch marks).



Results To identify the most parsimonious tree, we total all of the base-change events noted in steps 3 and 4. We conclude that the first tree is the most parsimonious of the three possible phylogenies. (In a real example, many more sites would be analyzed. Hence, the trees would often differ by more than one base-change event.)



a three-species problem. Computer programs use the principle of parsimony to estimate phylogenies in a similar way: They examine large numbers of possible trees and identify those that require the fewest evolutionary changes.

Phylogenetic Trees as Hypotheses

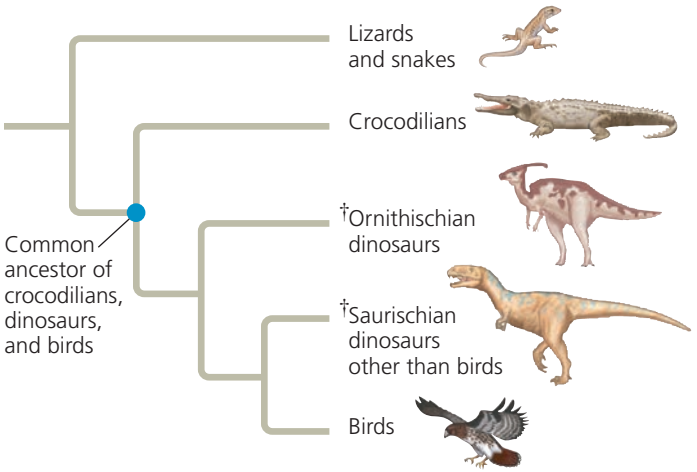
This is a good place to reiterate that any phylogenetic tree represents a hypothesis about how the organisms in the tree are related to one another. The best hypothesis is the one that best fits all the available data. A phylogenetic hypothesis may be modified when new evidence compels systematists to revise their trees. Indeed, while many older phylogenetic hypotheses have been supported by new morphological and molecular data, others have been changed or rejected.

Thinking of phylogenies as hypotheses also allows us to use them in a powerful way: We can make and test predictions based on the assumption that a particular phylogeny—our hypothesis—is correct. For example, in an approach known as *phylogenetic bracketing*, we can predict (by parsimony) that features shared by two groups of closely related organisms are present in their common ancestor and all of its descendants unless independent data indicate otherwise. (Note that "prediction" can refer to unknown past events as well as to evolutionary changes yet to occur.)

This approach has been used to make novel predictions about dinosaurs. For example, there is evidence that birds descended from the theropods, a group of bipedal saurischian dinosaurs. As seen in **Figure 22.16**, the closest living relatives of birds are crocodiles. Birds and crocodiles share numerous features: They have four-chambered hearts, they "sing" to defend territories and attract mates (although a crocodile's "song" is more like a bellow), and they build nests. Both birds and crocodiles also care for their eggs by *brooding*, a behavior in which a parent warms the eggs with its body. Birds brood by sitting on their eggs, whereas crocodiles cover their eggs with their neck. Reasoning that any feature shared by birds and crocodiles is likely to have been present in their common ancestor (denoted by the blue dot in Figure 22.16) and *all* of its descendants, biologists predicted that dinosaurs had four-chambered hearts, sang, built nests, and exhibited brooding.

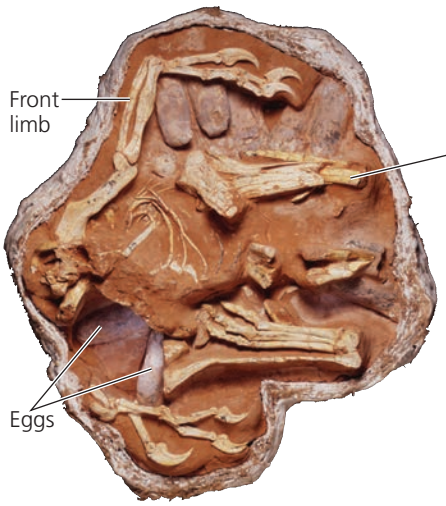
Internal organs, such as the heart, rarely fossilize, and it is, of course, difficult to test whether dinosaurs sang to defend territories and attract mates. However, fossilized dinosaur eggs and nests have provided evidence supporting the prediction of brooding in dinosaurs. First, a fossil embryo of an *Oviraptor* dinosaur was found, still inside its egg. This egg was identical to those found in another fossil, one that showed an *Oviraptor* crouching over a group of eggs in a posture similar to that seen in brooding birds today **(Figure 22.17)**. Researchers suggested that the *Oviraptor* dinosaur preserved in this second fossil died while incubating or protecting its eggs. The broader conclusion that emerged from this work—that dinosaurs built nests and exhibited brooding—has since been strengthened by additional fossil discoveries that show that other species of dinosaurs built nests and sat on their eggs. Finally, by supporting predictions based on the phylogenetic

Figure 22.16 A phylogenetic tree of birds and their close relatives. († indicates extinct lineages.)

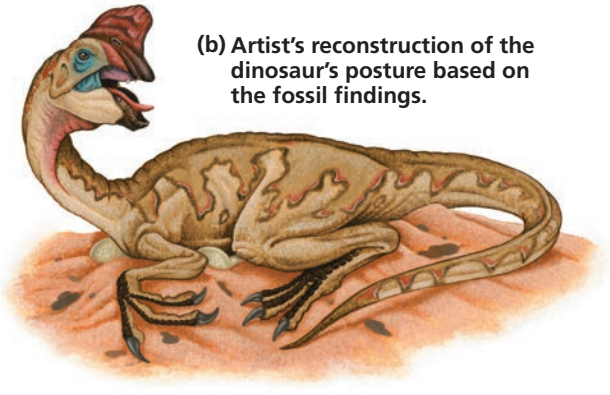


VISUAL SKILLS ➤ *In this tree diagram, what is the sister taxon of the clade that includes dinosaurs and their most recent common ancestor? Explain.*

Figure 22.17 Fossil support for a phylogenetic prediction: Dinosaurs built nests and brooded their eggs.



(a) Fossil remains of *Oviraptor* and eggs. The orientation of the bones, which surround and cover the eggs, suggests that the dinosaur died while incubating or protecting its eggs.



(b) Artist's reconstruction of the dinosaur's posture based on the fossil findings.

hypothesis shown in Figure 22.16, fossil discoveries of nests and brooding in dinosaurs provide independent data that suggest that the hypothesis is correct.

CONCEPT CHECK 22.3

1. To distinguish a particular clade of mammals within the larger clade that corresponds to class Mammalia, would hair be a useful character? Why or why not?
2. The most parsimonious tree of evolutionary relationships can be inaccurate. How can this occur?
3. **WHAT IF?** ➤ Draw a phylogenetic tree that includes the relationships from Figure 25.7 and Figure 22.16. Traditionally, all the taxa shown besides birds and mammals were classified as reptiles. Would a cladistic approach support that classification? Explain.

For suggested answers, see Appendix A.

CONCEPT 22.4

An organism's evolutionary history is documented in its genome

As you have seen in this chapter, comparisons of nucleic acids or other molecules can be used to deduce relatedness. In some cases, such comparisons can reveal phylogenetic relationships that cannot be determined by nonmolecular methods such as comparative anatomy. For example, the analysis of molecular data helps us uncover evolutionary relationships between groups that have little common ground for morphological comparison, such as animals and fungi. And molecular methods allow us to reconstruct phylogenies among groups of present-day organisms for which the fossil record is poor or lacking entirely.

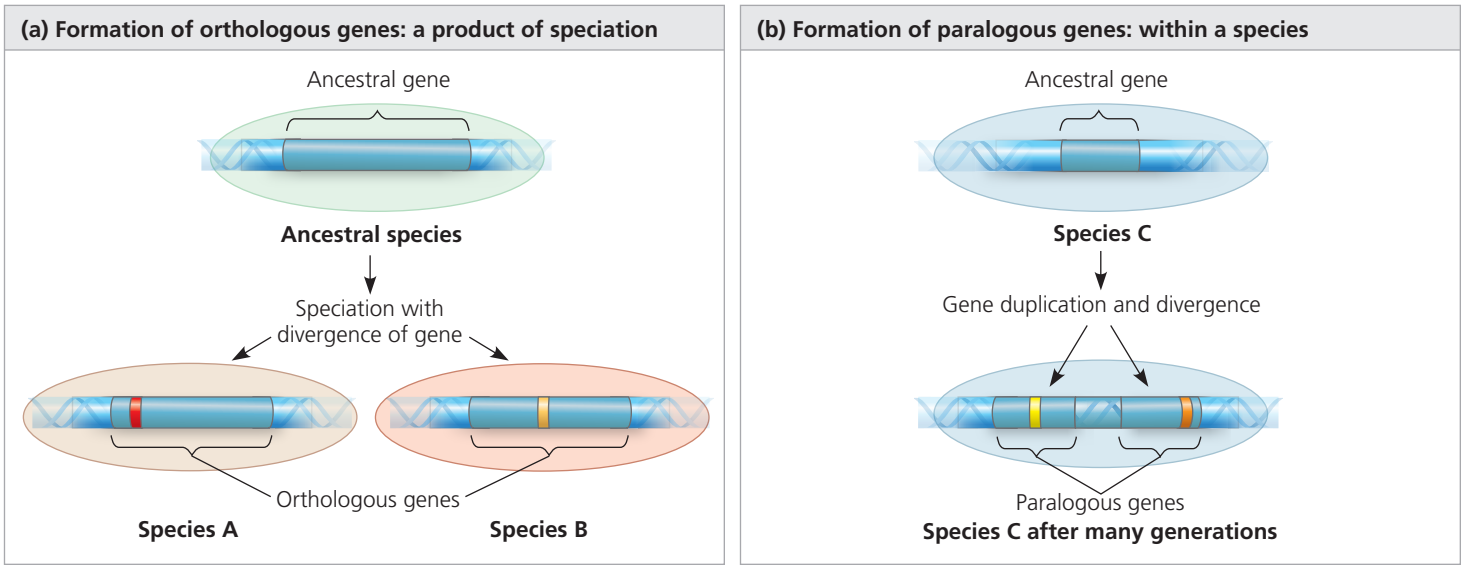
Different genes can evolve at different rates, even in the same evolutionary lineage. As a result, molecular trees can represent short or long periods of time, depending on which genes are used. For example, the DNA that codes for ribosomal RNA (rRNA) changes relatively slowly. Therefore, comparisons of DNA sequences in these genes are useful for investigating relationships between taxa that diverged hundreds of millions of years ago. Studies of rRNA sequences indicate, for instance, that fungi are more closely related to animals than to plants. In contrast, mitochondrial DNA (mtDNA) evolves relatively rapidly and can be used to explore recent evolutionary events. One research team has traced the relationships among Native American groups through their mtDNA sequences. The molecular findings corroborate other evidence that the Pima of Arizona, the Maya of Mexico, and the Yanomami of Venezuela are closely related, probably descending from the first of three waves of immigrants that crossed the Bering Land Bridge from Asia to the Americas about 15,000 years ago.

Gene Duplications and Gene Families

What do molecular data reveal about the evolutionary history of genome change? Consider gene duplication, which plays an important role in evolution because it increases the number of genes in the genome, providing more opportunities for further evolutionary changes. Molecular techniques now allow us to trace the phylogenies of gene duplications. These molecular phylogenies must account for repeated duplications that have resulted in *gene families*, groups of related genes within an organism's genome (see Figure 20.11).

Accounting for such duplications leads us to distinguish two types of homologous genes **(Figure 22.18)**: orthologous genes and paralogous genes. In **orthologous genes**

Figure 22.18 Two types of homologous genes. Colored bands mark regions of the genes where differences in base sequences have accumulated.



(from the Greek *orthos*, exact), the homology results from speciation (the splitting of one species into two or more species) and hence occurs between genes found in different species (see Figure 22.18a). For example, the genes that code for cytochrome *c* (a protein that functions in electron transport chains) in humans and dogs are orthologous. In **paralogous genes** (from the Greek *para*, in parallel), the homology results from gene duplication; hence, multiple copies of these genes have diverged from one another within a species (see Figure 22.18b). In Concept 23.1, you encountered the example of olfactory receptor genes, which have undergone many gene duplications in vertebrates; humans have 380 functional copies of these paralogous genes, while mice have 1,200.

Note that orthologous genes can diverge only after speciation has taken place, that is, after the genes are found in the genomes of different species. For example, although the cytochrome *c* genes in humans and dogs serve the same function, the gene's sequence in humans has diverged from that in dogs in the time since these species last shared a common ancestor. Paralogous genes, on the other hand, can diverge within a species because they are present in more than one copy in the genome. The paralogous genes that make up the olfactory receptor gene family in humans have diverged from each other during our long evolutionary history. They now specify proteins that confer sensitivity to a wide variety of molecules, ranging from food odors to sex pheromones.

Genome Evolution

Now that we can compare the entire genomes of different organisms, including our own, two patterns have emerged. First, lineages that diverged long ago often share many orthologous genes. For example, though the human and mouse lineages diverged about 65 million years ago, 99% of the genes of humans and mice are orthologous. And 50% of human genes are orthologous with those of yeast, despite 1 billion years of divergent evolution. Such commonalities explain why disparate organisms nevertheless share many biochemical and developmental pathways. As a result of these shared pathways, the functioning of genes linked to diseases in humans can often be investigated by studying yeast and other organisms distantly related to humans.

Second, the number of genes a species has doesn't seem to increase through duplication at the same rate as perceived phenotypic complexity. Humans have only about four times as many genes as yeast, a single-celled eukaryote, even though—unlike yeast—we have a large, complex brain and a body with more than 200 different types of tissues. Evidence is emerging that many human genes are more versatile than those of yeast: A single human gene can encode multiple proteins that perform different tasks in various body tissues. Unraveling the mechanisms that cause this genomic versatility and phenotypic variation is an exciting challenge.

CONCEPT CHECK 22.4

1. Explain how comparing proteins of two species can yield data about the species' evolutionary relationship.
2. **WHAT IF? ➤** Suppose gene A is orthologous in species 1 and species 2, and gene B is paralogous to gene A in species 1. Suggest a sequence of two evolutionary events that could result in the following: Gene A differs considerably between species, yet gene A and gene B show little divergence from each other.
3. **MAKE CONNECTIONS ➤** Review Figure 18.13, then suggest how a particular gene could have different functions in different tissues within an organism.

For suggested answers, see Appendix A.

CONCEPT 22.5

Molecular clocks help track evolutionary time

One goal of evolutionary biology is to understand the relationships among all organisms. It is also helpful to know when lineages diverged from one another, including those for which there is no fossil record. But how can we determine the timing of phylogenies that extend beyond the fossil record?

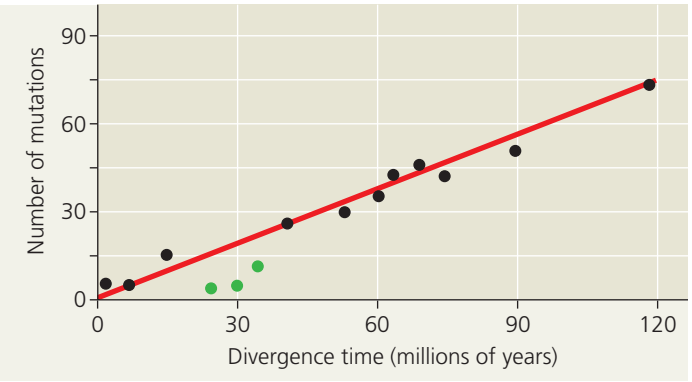
Molecular Clocks

We stated earlier that researchers have estimated that the common ancestor of Hawaiian silversword plants lived about 5 million years ago. How did they make this estimate? They relied on the concept of a **molecular clock**, an approach for measuring the absolute time of evolutionary change based on the observation that some genes and other regions of genomes appear to evolve at constant rates. An assumption underlying the molecular clock is that the number of nucleotide substitutions in orthologous genes is proportional to the time that has elapsed since the genes branched from their common ancestor. In the case of paralogous genes, the number of substitutions is proportional to the time since the ancestral gene was duplicated.

We can calibrate the molecular clock of a gene that has a reliable average rate of evolution by graphing the number of genetic differences—for example, nucleotide, codon, or amino acid differences—against the dates of evolutionary branch points that are known from the fossil record **(Figure 22.19)**. The average rates of genetic change inferred from such graphs can then be used to estimate the dates of events that cannot be discerned from the fossil record, such as the origin of the silverswords discussed earlier.

Of course, no gene marks time with complete precision. In fact, some portions of the genome appear to have evolved in irregular bursts that are not at all clocklike. And even those genes that seem to act as reliable molecular clocks

Figure 22.19 A molecular clock for mammals. The number of accumulated mutations in seven proteins has increased over time in a consistent manner for most mammal species. The three green data points represent primate species, whose proteins appear to have evolved more slowly than those of other mammals. The divergence time for each data point was based on fossil evidence.



INTERPRET THE DATA *Use the graph to estimate the divergence time for a mammal with a total of 30 mutations in the seven proteins.*

are accurate only in the statistical sense of showing a fairly smooth *average* rate of change. Over time, there may still be deviations from that average rate. Furthermore, the same gene may evolve at different rates in different groups of organisms. Finally, when comparing genes that are clocklike, the rate of the clock may vary greatly from one gene to another; some genes evolve a million times faster than others.

Differences in Clock Speed

What causes such differences in the speed at which clocklike genes evolve? The answer stems from the fact that some mutations are selectively neutral—neither beneficial nor detrimental. Of course, many new mutations are harmful and are removed quickly by selection. But if most of the rest are neutral and have little or no effect on fitness, then the rate of evolution of those neutral mutations should indeed be regular, like a clock. Differences in the clock rate for different genes are related to how important a gene is. If the exact sequence of amino acids that a gene specifies is essential to survival, most of the mutational changes will be harmful and only a few will be neutral. As a result, such genes change only slowly. But if the exact sequence of amino acids is less critical, fewer of the new mutations will be harmful and more will be neutral. Such genes change more quickly.

Potential Problems with Molecular Clocks

As we've seen, molecular clocks do not run as smoothly as would be expected if the underlying mutations were selectively neutral. Many irregularities are likely to be the result of natural selection in which certain DNA changes are favored over others. Indeed, evidence suggests that almost half the amino acid differences in proteins of two *Drosophila* species, *D. simulans* and *D. yakuba*, are not neutral but have resulted from natural selection. But because the direction of natural selection may change repeatedly over long periods of time (and hence may average out), some genes experiencing selection can nevertheless serve as approximate markers of elapsed time.

Another question arises when researchers attempt to extend molecular clocks beyond the time span documented by the fossil record. Although an abundant fossil record extends back only about 550 million years, molecular clocks have been used to date evolutionary divergences that occurred a billion or more years ago. These estimates assume that the clocks have been constant for all that time. Such estimates are highly uncertain.

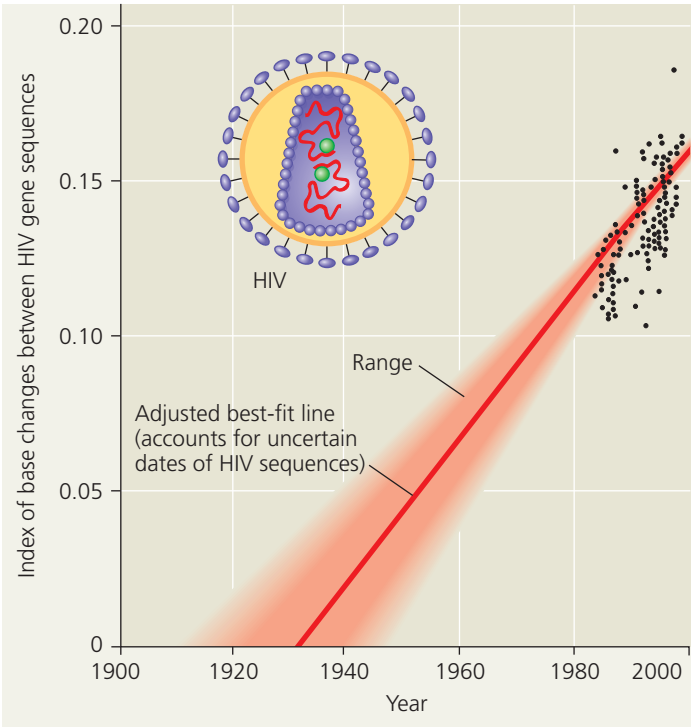
In some cases, problems may be avoided by calibrating molecular clocks with data on the rates at which genes have evolved in different taxa. In other cases, problems may be avoided by using many genes rather than just using one or a few genes. By using many genes, fluctuations in evolutionary rate due to natural selection or other factors that vary over time may average out. For example, one group of researchers constructed molecular clocks of vertebrate evolution from published sequence data for 658 nuclear genes. Despite the broad period of time covered (nearly 600 million years) and the fact that natural selection probably affected some of these genes, their estimates of divergence times agreed closely with fossil-based estimates. As this example suggests, if used with care, molecular clocks can aid our understanding of evolutionary relationships.

Applying a Molecular Clock: Dating the Origin of HIV

Researchers have used a molecular clock to date the origin of HIV infection in humans. Phylogenetic analysis shows that HIV, the virus that causes AIDS, is descended from viruses that infect chimpanzees and other primates. (Most of these viruses do not cause AIDS-like diseases in their native hosts.) When did HIV jump to humans? There is no simple answer because the virus has spread to humans more than once. The multiple origins of HIV are reflected in the variety of strains (genetic types) of the virus. HIV's genetic material is made of RNA, and like other RNA viruses, it evolves quickly.

The most widespread strain in humans is HIV-1 M. To pinpoint the earliest HIV-1 M infection, researchers compared samples of the virus from various times during the epidemic, including a sample from 1959. A comparison of gene sequences showed that the virus has evolved in a clocklike fashion. Extrapolating backward in time using the molecular

Figure 22.20 Dating the origin of HIV-1 M. The black data points are based on DNA sequences of an HIV gene in patients' blood samples. (The dates when these individual HIV gene sequences arose are not certain because a person can harbor the virus for years before symptoms occur.) Projecting the gene's rate of change backward in time by this method suggests that the virus originated in the 1930s.



clock indicates that the HIV-1 M strain first spread to humans around 1930 **(Figure 22.20)**. A later study, which dated the origin of HIV using a more advanced molecular clock approach than that covered in this book, estimated that the HIV-1 M strain first spread to humans around 1910.

CONCEPT CHECK 22.5

1. What is a molecular clock? What assumption underlies the use of a molecular clock?
2. **MAKE CONNECTIONS ➤** Review Concept 17.5. Then explain how numerous base changes could occur in an organism's DNA yet have no effect on its fitness.
3. **WHAT IF? ➤** Suppose a molecular clock dates the divergence of two taxa at 80 million years ago, but new fossil evidence shows that the taxa diverged at least 120 million years ago. Explain how this could happen.

For suggested answers, see Appendix A.

CONCEPT 22.6

Our understanding of the tree of life continues to change based on new data

The discovery that the glass lizard in Figure 22.1 evolved from a different lineage of legless lizards than did snakes is one example of how our understanding of life's diversity is informed by systematics. Indeed, in recent decades, systematists have gained insight into even the very deepest branches of the tree of life by analyzing DNA sequence data.

From Two Kingdoms to Three Domains

Taxonomists once classified all known species into two kingdoms: plants and animals. Classification schemes with more than two kingdoms gained broad acceptance in the late 1960s, when many biologists recognized five kingdoms: Monera (prokaryotes), Protista (a diverse kingdom consisting mostly of unicellular organisms), Plantae, Fungi, and Animalia. This system highlighted the two fundamentally different types of cells, prokaryotic and eukaryotic, and set the prokaryotes apart from all eukaryotes by placing them in their own kingdom, Monera.

However, phylogenies based on genetic data soon revealed a problem with this system: Some prokaryotes differ as much from each other as they do from eukaryotes. Such difficulties led biologists to adopt a three-domain system. The three domains—Bacteria, Archaea, and Eukarya—are a taxonomic level higher than the kingdom level. The validity of these domains has been supported by many studies, including a recent study that analyzed nearly 100 completely sequenced genomes.

The domain Bacteria contains most of the currently known prokaryotes, while the domain Archaea consists of a diverse group of prokaryotic organisms that inhabit a wide variety of environments. The domain Eukarya consists of all the organisms that have cells containing true nuclei. This domain includes many groups of single-celled organisms as well as multicellular plants, fungi, and animals. **Figure 22.21** represents one possible phylogenetic tree for the three domains and some of the many lineages they encompass.

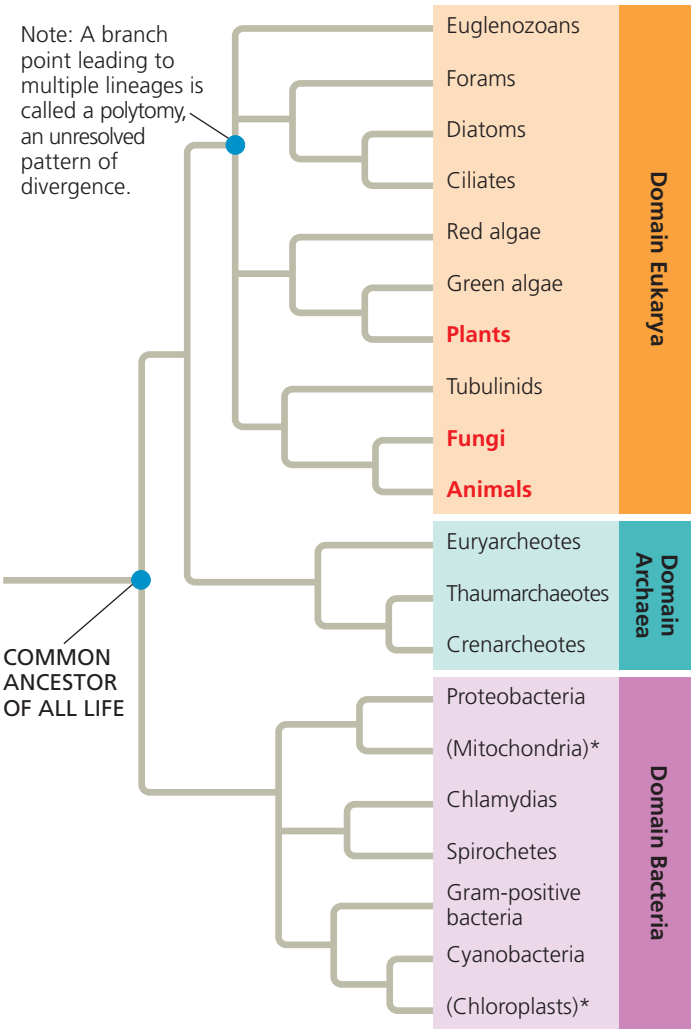
The three-domain system highlights the fact that much of the history of life has been about single-celled organisms. The two prokaryotic domains consist entirely of single-celled organisms, and even in Eukarya, only the branches labeled in red type (plants, fungi, and animals) are dominated by multicellular organisms. Of the five kingdoms previously recognized by taxonomists, most biologists continue to recognize Plantae, Fungi, and Animalia, but not Monera and Protista. The kingdom Monera is obsolete because it would have members in two different domains. The kingdom Protista has also crumbled because it includes members that are more closely related to plants, fungi, or animals than to other protists (see Figure 28.2).

New research continues to change our understanding of the tree of life. For example, in the past decade, metagenomic studies have uncovered the genomes of many new species of archaea, leading to discovery of the Thaumarchaeota and other previously unknown phyla of archaea (see Concept 27.4).

The Important Role of Horizontal Gene Transfer

In the phylogeny shown in Figure 22.21, the first major split in the history of life occurred when bacteria diverged from

Figure 22.21 The three domains of life. This phylogenetic tree is based on sequence data for rRNA and other genes. For simplicity, only some of the major branches in each domain are shown. Lineages within Eukarya that are dominated by multicellular organisms (plants, fungi, and animals) are in red type, while the two lineages denoted by an asterisk are based on DNA from cellular organelles. All other lineages consist solely or mainly of single-celled organisms.



MAKE CONNECTIONS ➤ *After reviewing endosymbiont theory (see Figure 7.16), explain the specific positions of the mitochondrion and chloroplast lineages on this tree.*

other organisms. If this tree is correct, eukaryotes and archaea are more closely related to each other than either is to bacteria.

This reconstruction of the tree of life is based in part on sequence comparisons of rRNA genes, which code for the RNA components of ribosomes. However, some other genes reveal a different set of relationships. For example, researchers have found that many of the genes that influence metabolism in yeast (a unicellular eukaryote) are more similar to genes in the domain Bacteria than they are to genes in the domain Archaea—a finding that suggests that the eukaryotes may share a more recent common ancestor with bacteria than with archaea.

What causes trees based on data from different genes to yield such different results? Comparisons of complete genomes from the three domains show that there have been substantial movements of genes between organisms in the different domains. These took place through **horizontal gene transfer**, a process in which genes are transferred from one genome to another through mechanisms such as exchange of transposable elements and plasmids, viral infection (see Concept 26.2), and perhaps fusions of organisms (as when a host and its endosymbiont become a single organism). Recent research reinforces the view that horizontal gene transfer is important. For example, one study found that on average, 80% of the genes in 181 prokaryotic genomes had moved between species at some point during the course of evolution. Because phylogenetic trees are based on the assumption that genes are passed vertically from one generation to the next, the occurrence of such horizontal transfer events helps to explain why trees built using different genes can give inconsistent results.

Horizontal gene transfer can also occur between eukaryotes. For example, over 200 cases of the horizontal transfer of transposons have been reported in eukaryotes, including humans and other primates, plants, birds, and lizards. Nuclear genes have also been transferred horizontally from one eukaryote to another. The **Scientific Skills Exercise** describes one such example, giving you the opportunity to interpret data collected by Nancy Moran on the transfer of a pigment gene to an aphid from another species.

Recent evidence indicates that eukaryotes can even acquire nuclear genes from bacteria and archaea. For example, a 2013 genomic analysis showed that the alga *Galdieria sulphuraria* **(Figure 22.22)** acquired about 5% of its genes from various bacterial and archaeal species. Unlike most eukaryotes, this alga can survive in environments that are highly acidic or extremely hot, as well as those with high concentrations of heavy metals. The researchers identified specific genes transferred from prokaryotes that have enabled *G. sulphuraria* to thrive in such extreme habitats.

Overall, horizontal gene transfer has played a key role throughout the evolutionary history of life, and it continues to occur today. Some biologists have argued that horizontal gene transfer was so common that the early history of life should be represented not as a dichotomously branching tree

Figure 22.22 A recipient of transferred genes: the alga *Galdieria sulphuraria*. Genes received from prokaryotes enable *G. sulphuraria* (inset) to grow in extreme environments, including on sulfur-encrusted rocks around volcanic hot springs similar to this one in Yellowstone National Park.

SCIENTIFIC SKILLS EXERCISE

Using Protein Sequence Data to Test an Evolutionary Hypothesis

Did Aphids Acquire Their Ability to Make Carotenoids Through Horizontal Gene Transfer? Carotenoids are colored molecules that have diverse functions in many organisms, such as photosynthesis in plants and light detection in animals. Plants and many microorganisms can synthesize carotenoids from scratch, but animals generally cannot (they must obtain carotenoids from their diet). One exception is the pea aphid *Acyrthosiphon pisum*, a small, plant-dwelling insect whose genome includes a full set of genes for the enzymes needed to make carotenoids. Because other animals lack these genes, it is unlikely that aphids inherited them from a single-celled common ancestor shared with microorganisms and plants. So where did they come from? Evolutionary biologists hypothesized that an aphid ancestor acquired these genes by horizontal gene transfer from distantly related organisms.

How the Experiment Was Done Scientists obtained the DNA sequences for the carotenoid-biosynthesis genes from several species, including aphids, fungi, bacteria, and plants. A computer "translated" these sequences into amino acid sequences of the encoded polypeptides and aligned the amino acid sequences. This allowed the team to compare the corresponding polypeptides in the different organisms.

Data from the Experiment The sequences below show the first 60 amino acids of one polypeptide of the carotenoid-biosynthesis enzymes in the plant *Arabidopsis thaliana* (bottom) and the corresponding amino acids in five nonplant species, using the one-letter abbreviations for the amino acids (see Figure 5.14). A dash (–) indicates a gap inserted in a sequence to optimize its alignment with the corresponding sequence in *Arabidopsis*.

INTERPRET THE DATA

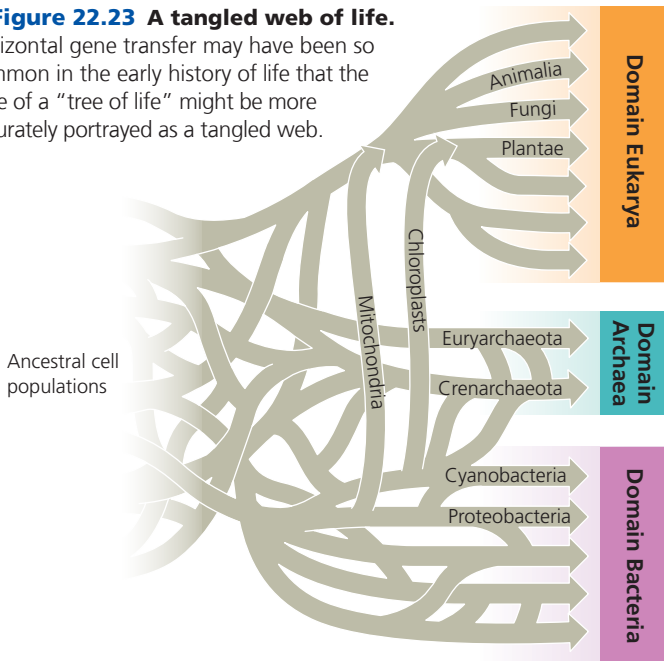
1. In the rows of data for the organisms being compared with the aphid, highlight the amino acids that are identical to the corresponding amino acids in the aphid.
2. Which organism has the most amino acids in common with the aphid? Rank the partial polypeptides from the other four organisms in degree of similarity to that of the aphid.
3. Do these data support the hypothesis that aphids acquired the gene for this polypeptide by horizontal gene transfer? Why or why not? If horizontal gene transfer did occur, what type of organism is likely to have been the source?
4. What additional sequence data would support your hypothesis?
5. How would you account for the similarities between the aphid sequence and the sequences for the bacteria and plant?

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Data from Nancy A. Moran, Yale University. See N. A. Moran and T. Jarvik, Lateral transfer of genes from fungi underlies carotenoid production in aphids, *Science* 328:624–627 (2010).

Organism	Alignment of Amino Acid Sequences
Acyrthosiphon (aphid)	IKIIIIGSGV GGTAAAARLS KKGFQVEVYE KNSYNGGRCS IIR-HNGHRF DQGPSL--YL
Ustilago (fungus)	KKVVIIGAGA GGTALAARLG RRGYSVTVLE KNSFGGGRCS LIH-HDGHRW DQGPSL--YL
Gibberella (fungus)	KSVIVIGAGV GGVSTAARLA KAGFKVTILE KNDFTGGRCS LIH-NDGHRF DQGPSL--LL
Staphylococcus (bacterium)	MKIAVIGAGV TGLAAAARIA SQGHEVTIFE KNNNVGGRMN QLK-KDGFTF DMGPTI--VM
Pantoea (bacterium)	KRTFVIGAGF GGLALAIRLQ AAGIATTVLE QHDKPGGRAY VWQ-DQGFTF DAGPTV--IT
Arabidopsis (plant)	WDAVVIGGGH NGLTAAAYLA RGGLSVAVLE RRHVIGGAAV TEEIVPGFKF SRCSYLQGLL

➤ Figure 22.23 A tangled web of life. Horizontal gene transfer may have been so common in the early history of life that the base of a "tree of life" might be more accurately portrayed as a tangled web.



like that in Figure 22.21, but rather as a tangled network of connected branches **(Figure 22.23)**. Although scientists continue to debate how best to portray the earliest steps in the history of life, in recent decades there have been many exciting discoveries about evolutionary events that occurred over time. We'll explore the mechanisms that underlie such events in the rest of this unit's chapters, beginning with the evolution of populations.

CONCEPT CHECK 22.6

1. Why is the kingdom Monera no longer considered a valid taxon?
2. Explain why phylogenies based on different genes can yield different branching patterns for the tree of all life.
3. **MAKE CONNECTIONS ➤** Explain how the origin of eukaryotes is thought to have represented a fusion of organisms, leading to extensive horizontal gene transfer. (See Figure 25.10.)

For suggested answers, see Appendix A.

22 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

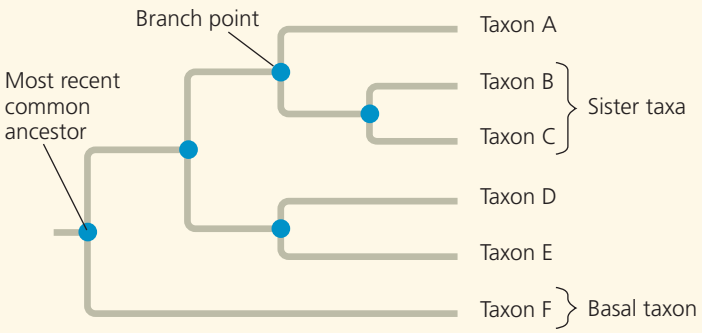
SUMMARY OF KEY CONCEPTS

CONCEPT 22.1

Phylogenies show evolutionary relationships *(pp. 520–524)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

- Linnaeus's **binomial** classification system gives organisms two-part names: a **genus** plus a specific epithet.
- In the Linnaean system, species are grouped in increasingly broad **taxa**: Related genera are placed in the same **family**, families in **orders**, orders in **classes**, classes in **phyla**, phyla in **kingdoms**, and (more recently) kingdoms in **domains**.
- Systematists depict evolutionary relationships as branching **phylogenetic trees**. Many systematists propose that classification be based entirely on evolutionary relationships.



- Unless branch lengths are proportional to time or genetic change, a phylogenetic tree indicates only patterns of descent.
- Much information can be learned about a species from its evolutionary history; hence, phylogenies are useful in a wide range of applications.

? *Humans and chimpanzees are sister species. Explain what this statement means.*

CONCEPT 22.2

Phylogenies are inferred from morphological and molecular data *(pp. 524–525)*

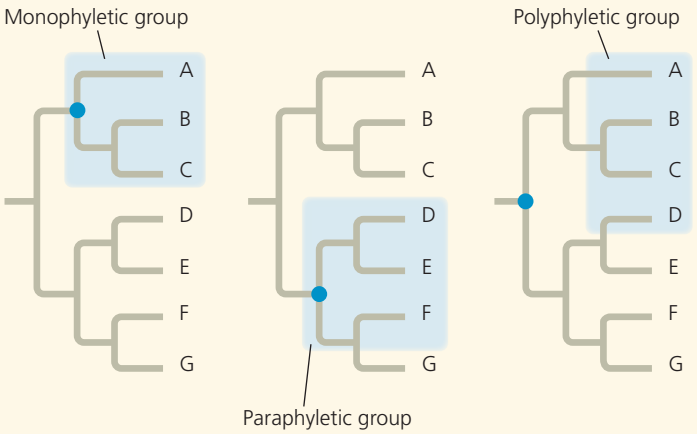
- Organisms with similar morphologies or DNA sequences are likely to be more closely related than organisms with very different structures and genetic sequences.
- To infer phylogeny, **homology** (similarity due to shared ancestry) must be distinguished from **analogy** (similarity due to convergent evolution).
- Computer programs are used to align comparable DNA sequences and to distinguish molecular homologies from coincidental matches between taxa that diverged long ago.

? *Why is it necessary to distinguish homology from analogy to infer phylogeny?*

CONCEPT 22.3

Shared characters are used to construct phylogenetic trees *(pp. 525–531)*

- A **clade** is a monophyletic group that includes an ancestral species and all of its descendants.
- Clades can be distinguished by their **shared derived characters**.



- Among phylogenies, the most parsimonious tree is the one that requires the fewest evolutionary changes. The most likely tree is the one based on the most likely pattern of changes.
- Well-supported phylogenetic hypotheses are consistent with a wide range of data.

? *Explain the logic of using shared derived characters to infer phylogeny.*

CONCEPT 22.4

An organism's evolutionary history is documented in its genome *(pp. 531–532)*

- **Orthologous genes** are homologous genes found in different species as a result of speciation. **Paralogous genes** are homologous genes within a species that result from gene duplication; such genes can diverge and potentially take on new functions.
- Distantly related species often have many orthologous genes. The small variation in gene number in organisms of varying complexity suggests that genes are versatile and may have multiple functions.

? *When reconstructing phylogenies, is it more useful to compare orthologous or paralogous genes? Explain.*

CONCEPT 22.5

Molecular clocks help track evolutionary time *(pp. 532–534)*

- Some regions of DNA change at a rate consistent enough to serve as a **molecular clock**, a method of estimating the date of past evolutionary events based on the amount of genetic change. Other DNA regions change in a less predictable way.
- Molecular clock analyses suggest that the most common strain of HIV jumped from primates to humans in the early 1900s.

? *Describe some assumptions and limitations of molecular clocks.*

CONCEPT 22.6

Our understanding of the tree of life continues to change based on new data *(pp. 534–536)*

- Past classification systems have given way to the current view of the tree of life, which consists of three great domains: Bacteria, Archaea, and Eukarya.

- Phylogenies based in part on rRNA genes suggest that eukaryotes are most closely related to archaea, while data from some other genes suggest a closer relationship to bacteria.
- Genetic analyses indicate that extensive **horizontal gene transfer** has occurred throughout the evolutionary history of life.

? *Why was the five-kingdom system abandoned for a three-domain system?*

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–7 can be found in the Study Area in MasteringBiology.

8. EVOLUTION CONNECTION Darwin suggested looking at a species' close relatives to learn what its ancestors may have been like. Explain how his suggestion anticipates recent methods, such as phylogenetic bracketing and the use of outgroups in cladistic analysis.

PRACTICE TEST
goo.gl/iAsVgL

9. SCIENTIFIC INQUIRY • DRAW IT (a) Draw a phylogenetic tree based on characters 1–5 in the table below. Place hatch marks on the tree to indicate the origin(s) of characters 1–6. (b) Assume that tuna and dolphins are sister species and redraw the phylogenetic tree accordingly. Use hatch marks to indicate the origin(s) of characters 1–6. (c) Determine how many evolutionary changes are required in each tree. Identify the most parsimonious tree.

Character	Lancelet (outgroup)	Lamprey	Tuna	Salamander	Turtle	Leopard	Dolphin
(1) Backbone	0	1	1	1	1	1	1
(2) Hinged jaw	0	0	1	1	1	1	1
(3) Four limbs	0	0	0	1	1	1	1*
(4) Amnion	0	0	0	0	1	1	1
(5) Milk	0	0	0	0	0	1	1
(6) Dorsal fin	0	0	1	0	0	0	1

*Although adult dolphins have only two obvious limbs (their flippers), as embryos they have two hind-limb buds, for a total of four limbs.

10. WRITE ABOUT A THEME: INFORMATION In a short essay (100–150 words), explain how genetic information—along with an understanding of the process of descent with modification—enables scientists to reconstruct phylogenies that extend hundreds of millions of years back in time.

11. SYNTHESIZE YOUR KNOWLEDGE

This West Indian manatee (*Trichechus manatus*) is an aquatic mammal. Like amphibians and reptiles, mammals are tetrapods (vertebrates with four limbs). Explain why manatees are considered tetrapods even though they lack hind limbs, and suggest traits that manatees likely share with leopards and other mammals (see Figure 22.12b). Discuss how early members of the manatee lineage might have differed from today's manatees.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

This page is intentionally left blank.

23 Microevolution

Figure 23.1 Is this finch evolving?

KEY CONCEPTS

23.1 Genetic variation makes evolution possible

23.2 The Hardy-Weinberg equation can be used to test whether a population is evolving

23.3 Natural selection, genetic drift, and gene flow can alter allele frequencies in a population

23.4 Natural selection is the only mechanism that consistently causes adaptive evolution

The Smallest Unit of Evolution

One common misconception about evolution is that individual organisms evolve. It is true that natural selection acts on individuals: Each organism's traits affect its survival and reproductive success compared with those of other individuals. But the evolutionary impact of natural selection is only apparent in how a *population* of organisms changes over time.

Consider the medium ground finch (*Geospiza fortis*), a seed-eating bird that inhabits the Galápagos Islands **(Figure 23.1)**. In 1977, the *G. fortis* population on the island of Daphne Major was decimated by a long period of drought: Of some 1,200 birds, only 180 survived. Researchers Peter and Rosemary Grant observed that during the drought, small, soft seeds were in short supply. The finches mostly fed on large, hard seeds that were more plentiful. Birds with larger, deeper beaks were better able to crack and eat these larger seeds, and they survived at a higher rate than did finches with smaller beaks. Since beak depth is an inherited trait in these birds, the offspring of surviving birds also tended to have deep beaks. As a result, the average beak depth in the next generation of *G. fortis* was greater than it had been in the pre-drought population **(Figure 23.2)**. The finch population had evolved by natural selection. However, the *individual* finches did not evolve. Each bird had a beak of a particular size, which did not grow larger during the drought. Rather, the proportion of large

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

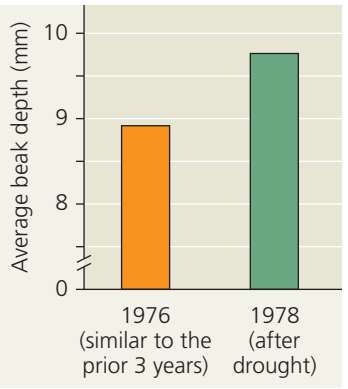


Figure 23.2 Evidence of selection by food source. The data represent adult beak depth measurements of medium ground finches hatched in the generations before and after the 1977 drought. In one generation, natural selection resulted in a larger average beak size in the population.

Instructors: A related Experimental Inquiry Tutorial can be assigned in MasteringBiology.

beaks in the population increased from generation to generation: The population evolved, not its individual members.

Focusing on evolutionary change in populations, we can define evolution on its smallest scale, called **microevolution**, as a change in allele frequencies in a population over generations. As you will see in this chapter, natural selection is not the only cause of microevolution. In fact, there are three main mechanisms that can cause allele frequency change: natural selection, genetic drift (chance events that alter allele frequencies), and gene flow (the transfer of alleles between populations). Each of these mechanisms has distinctive effects on the genetic composition of populations. However, only natural selection consistently improves the degree to which organisms are well suited for life in their environment (adaptation). Before we examine natural selection and adaptation more closely, let's revisit a prerequisite for these processes in a population: genetic variation.

CONCEPT 23.1

Genetic variation makes evolution possible

In *The Origin of Species*, Darwin provided abundant evidence that life on Earth has evolved over time, and he proposed natural selection as the primary mechanism for that change. He observed that individuals differ in their inherited traits and that selection acts on such differences, leading to evolutionary change. Although Darwin realized that variation in heritable traits is a prerequisite for evolution, he did not know precisely how organisms pass heritable traits to their offspring.

Just a few years after Darwin published *The Origin of Species*, Gregor Mendel wrote a groundbreaking paper on inheritance in pea plants (see Concept 14.1). In that paper, Mendel proposed a model of inheritance in which organisms transmit discrete heritable units (now called genes) to their offspring. Although Darwin did not know about genes, Mendel's paper set the stage for understanding the genetic differences on

Figure 23.3 Phenotypic variation in horses. In horses, coat color varies along a continuum and is influenced by multiple genes.

which evolution is based. Here we'll examine such genetic differences and how they are produced.

Genetic Variation

Individuals within all species vary in their phenotypic traits. Among humans, for example, you can easily observe phenotypic variation in facial features, height, and voice. Indeed, individual variation occurs in all species. And though you cannot identify a person's blood group (A, B, AB, or O) from his or her appearance, this and many other molecular traits also vary extensively among individuals.

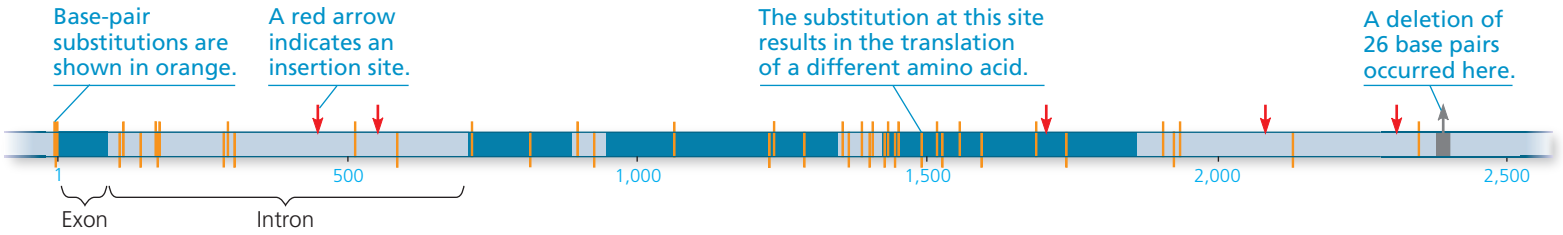
Such phenotypic variations often reflect **genetic variation**, differences among individuals in the composition of their genes or other DNA sequences. Some heritable phenotypic differences occur on an "either-or" basis, such as the flower colors of Mendel's pea plants: Each plant had flowers that were either purple or white (see Figure 14.3). Characters that vary in this way are typically determined by a single gene locus, with different alleles producing distinct phenotypes. In contrast, other phenotypic differences vary in gradations along a continuum. Such variation usually results from the influence of two or more genes on a single phenotypic character. In fact, many phenotypic characters are influenced by multiple genes, including coat color in horses **(Figure 23.3)**, seed number in maize (corn), and height in humans.

How much do genes and other DNA sequences vary from one individual to another? Genetic variation at the whole-gene level (*gene variability*) can be quantified as the average percentage of loci that are heterozygous. (Recall that a heterozygous individual has two different alleles for a given locus, whereas a homozygous individual has two identical alleles for that locus.) As an example, on average the fruit fly *Drosophila melanogaster* is heterozygous for about 1,920 of its 13,700 loci (14%) and homozygous for all the rest.

Considerable genetic variation can also be measured at the molecular level of DNA (*nucleotide variability*). But little of this variation results in phenotypic variation. Why? Many nucleotide variations occur within *introns*, noncoding

Figure 23.4 Extensive genetic variation at the molecular level. This diagram summarizes data from a study comparing the DNA sequence of the alcohol dehydrogenase (*Adh*) gene in several fruit flies (*Drosophila melanogaster*). The *Adh* gene has four exons (dark blue) separated by introns (light blue); the exons include the coding regions that are ultimately translated into the amino acids of the *Adh* enzyme (see Figure 5.1). Only one substitution has a phenotypic effect, producing a different form of the *Adh* enzyme.

MAKE CONNECTIONS ➤ *Review Figures 17.6 and 17.11. Explain how a base-pair substitution that alters a coding region of the* Adh *locus could have no effect on amino acid sequence. Then explain how an insertion in an exon could have no effect on the protein produced.*



segments of DNA lying between *exons*, the regions retained in mRNA after RNA processing (see Figure 17.12). And of the variations that occur within exons, most do not cause a change in the amino acid sequence of the protein encoded by the gene. For example, in the sequence comparison shown in **Figure 23.4**, there are 43 nucleotide sites with variable base pairs (where substitutions have occurred), as well as several sites where insertions or deletions have occurred. Although 18 variable sites occur within the four exons of the *Adh* gene, only one of these variations (at site 1,490) results in an amino acid change. Note, however, that this single variable site is enough to cause genetic variation at the level of the gene—and hence two different forms of the *Adh* enzyme are produced.

It is important to bear in mind that some phenotypic variation does not result from genetic differences among individuals (**Figure 23.5** shows a striking example in a caterpillar of the southwestern United States). Phenotype is the product of an inherited genotype and many environmental influences (see Concept 14.3). In a human example, bodybuilders alter their phenotypes dramatically but do not pass their huge muscles on to the next generation. In general, only the genetically determined part of phenotypic variation can have evolutionary consequences. As such, genetic variation provides the raw material for evolutionary change: Without genetic variation, evolution cannot occur.

Sources of Genetic Variation

The genetic variation on which evolution depends originates when mutation, gene duplication, or other processes produce new alleles and new genes. Genetic variants can be produced rapidly in organisms with short generation times. Sexual reproduction can also result in genetic variation as existing genes are arranged in new ways.

Formation of New Alleles

New alleles can arise by *mutation*, a change in the nucleotide sequence of an organism's DNA. A change of as little as one base in a gene—a "point mutation"—can have a significant impact on phenotype, as in sickle-cell disease (see Figure 17.26). We might expect that this would be the case: Organisms reflect many generations of past selection, and hence their phenotypes tend to be suited for life in their environments. As a result, most new mutations that alter a phenotype are at least slightly harmful.

In some cases, natural selection quickly removes such harmful alleles. In diploid organisms, however, harmful alleles that are recessive can be hidden from selection. Indeed, a harmful recessive allele can persist for generations by propagation in heterozygous individuals (where its harmful effects can be masked by the more favorable dominant allele). Such "heterozygote protection" maintains a huge pool of alleles that might not be favored under present conditions, but that could be beneficial if the environment changes.

While many mutations are harmful, many others are not. Recall that much of the DNA in eukaryotic genomes does not encode proteins (see Figure 20.6). Point mutations in these noncoding regions generally result in **neutral variation**, differences in DNA sequence that do not confer a selective advantage or disadvantage. The redundancy in the genetic code is another source of neutral variation: Even a point mutation in a gene that encodes a protein

Figure 23.5 Nonheritable variation. These caterpillars of the moth *Nemoria arizonaria* owe their different appearances to chemicals in their diets, not to differences in their genotypes. **(a)** Caterpillars raised on a diet of oak flowers resemble the flowers, whereas **(b)** their siblings raised on oak leaves resemble oak twigs.

will have no effect on the protein's function if the amino acid composition is not changed. And even where there is a change in the amino acid, it may not affect the protein's shape and function. Moreover, as you will see later in this chapter, a mutant allele may on rare occasions actually make its bearer better suited to the environment, enhancing reproductive success.

Finally, note that in multicellular organisms, only mutations in cell lines that produce gametes can be passed to offspring. In plants and fungi, this is not as limiting as it may sound, since many different cell lines can produce gametes. But in most animals, the majority of mutations occur in somatic cells and are not passed to offspring.

Altering Gene Number or Position

Chromosomal changes that delete, disrupt, or rearrange many loci are usually harmful. However, when such large-scale changes leave genes intact, they may not affect the organisms' phenotype. In rare cases, chromosomal rearrangements may even be beneficial. For example, the translocation of part of one chromosome to a different chromosome could link genes in a way that produces a positive effect.

A key potential source of variation is the duplication of genes due to errors in meiosis (such as unequal crossing over), slippage during DNA replication, or the activities of transposable elements (see Concept 20.5). Duplications of large chromosome segments, like other chromosomal aberrations, are often harmful, but the duplication of smaller pieces of DNA may not be. Gene duplications that do not have severe effects can persist over generations, allowing mutations to accumulate. The result is an expanded genome with new genes that may take on new functions.

Such increases in gene number appear to have played a major role in evolution. For example, the remote ancestors of mammals had a single gene for detecting odors that has since been duplicated many times. As a result, humans today have about 380 functional olfactory receptor genes, and mice have about 1,200. This dramatic proliferation of olfactory genes probably helped early mammals, enabling them to detect faint odors and to distinguish among many different smells.

Rapid Reproduction

Mutation rates tend to be low in plants and animals, averaging about one mutation in every 100,000 genes per generation, and they are often even lower in prokaryotes. But prokaryotes have many more generations per unit of time, so mutations can quickly generate genetic variation in their populations. The same is true of viruses. For instance, HIV has a generation time of about two days (that is, it takes two days for a newly formed virus to produce the next generation of viruses). HIV also has an RNA genome, which has a much higher mutation rate than a typical DNA genome because of the lack of RNA repair mechanisms in host cells (see Concept 26.2). For this reason, single-drug treatments are unlikely to be effective against HIV: Mutant forms of the virus that are resistant to a particular drug would tend to proliferate in relatively short order. The most effective AIDS treatments to date have been drug "cocktails" that combine several medications. This approach has worked well because it is less likely that a set of mutations that together confer resistance to *all* the drugs will occur in a short time period.

Sexual Reproduction

In organisms that reproduce sexually, most of the genetic variation in a population results from the unique combination of alleles that each individual receives from its parents. Of course, at the nucleotide level, all the differences among these alleles have originated from past mutations. Sexual reproduction then shuffles existing alleles and deals them at random to produce individual genotypes.

Three mechanisms contribute to this shuffling: crossing over, independent assortment of chromosomes, and fertilization (see Concept 13.4). During meiosis, homologous chromosomes, one inherited from each parent, trade some of their alleles by crossing over. These homologous chromosomes and the alleles they carry are then distributed at random into gametes. Then, because myriad possible mating combinations exist in a population, fertilization typically brings together gametes that have different genetic backgrounds. The combined effects of these three mechanisms ensure that sexual reproduction rearranges existing alleles into fresh combinations each generation, providing much of the genetic variation that makes evolution possible.

Animation: Origins of Genetic Variation

CONCEPT CHECK 23.1

1. Explain why genetic variation within a population is a prerequisite for evolution.
2. Of all the mutations that occur in a population, why do only a small fraction become widespread?
3. **MAKE CONNECTIONS >** If a population stopped reproducing sexually (but still reproduced asexually), how would its genetic variation be affected over time? Explain. (See Concept 13.4.)

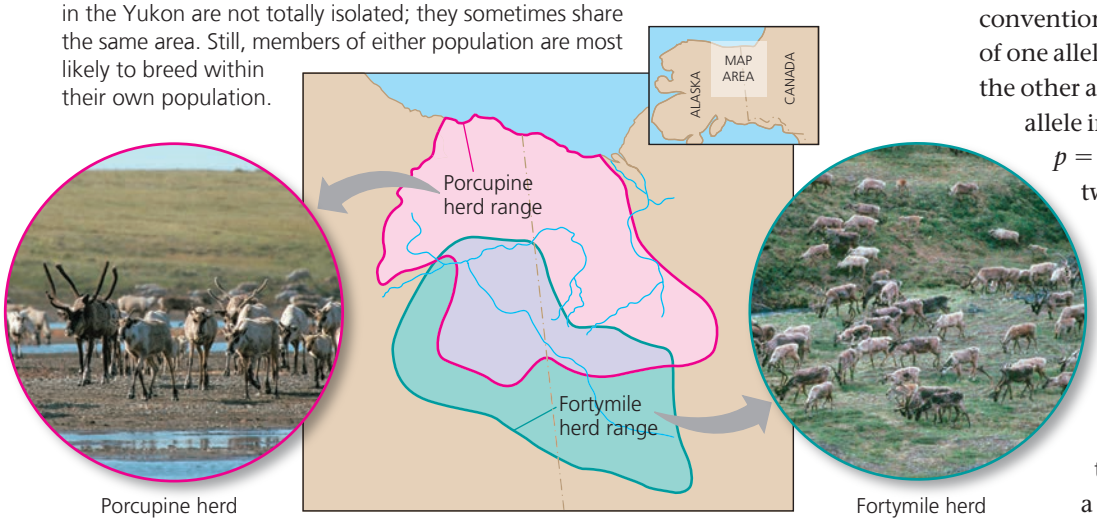
For suggested answers, see Appendix A.

CONCEPT 23.2

The Hardy-Weinberg equation can be used to test whether a population is evolving

Although the individuals in a population must differ genetically for evolution to occur, the presence of genetic variation does not guarantee that a population will evolve. For that to happen, one or more factors that cause evolution must be at work. In this section, we'll explore one way to test whether evolution is occurring in a population. First, let's clarify what we mean by a population.

Figure 23.6 One species, two populations. These two caribou populations in the Yukon are not totally isolated; they sometimes share the same area. Still, members of either population are most likely to breed within their own population.



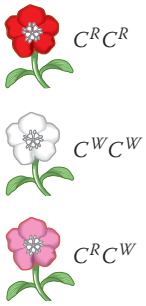
When studying a locus with two alleles, the convention is to use p to represent the frequency of one allele and q to represent the frequency of the other allele. Thus, p, the frequency of the C^R allele in the gene pool of this population, is $p = 0.8$ (80%). And because there are only two alleles for this gene, the frequency of the C^W allele, represented by q, must be $q = 1 - p = 0.2$ (20%). For loci that have more than two alleles, the sum of all allele frequencies must still equal 1 (100%).

Next we'll see how allele and genotype frequencies can be used to test whether evolution is occurring in a population.

Gene Pools and Allele Frequencies

A **population** is a group of individuals of the same species that live in the same area and interbreed, producing fertile offspring. Different populations of a species may be isolated geographically from one another, exchanging genetic material only rarely. Such isolation is common for species that live on widely separated islands or in different lakes. But not all populations are isolated **(Figure 23.6)**. Still, members of a population typically breed with one another and thus on average are more closely related to each other than to members of other populations.

We can characterize a population's genetic makeup by describing its **gene pool**, which consists of all copies of every type of allele at every locus in all members of the population. If only one allele exists for a particular locus in a population, that allele is said to be *fixed* in the gene pool, and all individuals are homozygous for that allele. But if there are two or more alleles for a particular locus in a population, individuals may be either homozygous or heterozygous.



For example, imagine a population of 500 wildflower plants with two alleles, C^R and C^W, for a locus that codes for flower pigment. These alleles show incomplete dominance; thus, each genotype has a distinct phenotype. Plants homozygous for the C^R allele (C^RC^R) produce red pigment and have red flowers; plants homozygous for the C^W allele (C^WC^W) produce no red pigment and have white flowers; and heterozygotes (C^RC^W) produce some red pigment and have pink flowers.

Each allele has a frequency (proportion) in the population. For example, suppose our population has 320 plants with red flowers, 160 with pink flowers, and 20 with white flowers. Because these are diploid organisms, these 500 individuals have a total of 1,000 copies of the gene for flower color. The C^R allele accounts for 800 of these copies ($320 \times 2 = 640$ for C^RC^R plants, plus $160 \times 1 = 160$ for C^RC^W plants). Thus, the frequency of the C^R allele is 800/1,000 = 0.8 (80%).

The Hardy-Weinberg Equation

One way to assess whether natural selection or other factors are causing evolution at a particular locus is to determine what the genetic makeup of a population would be if it were *not* evolving at that locus. We can then compare that scenario with the data we actually observed for the population. If there are no differences, we can conclude that the population is not evolving. If there are differences, this suggests that the population may be evolving—and then we can try to figure out why.

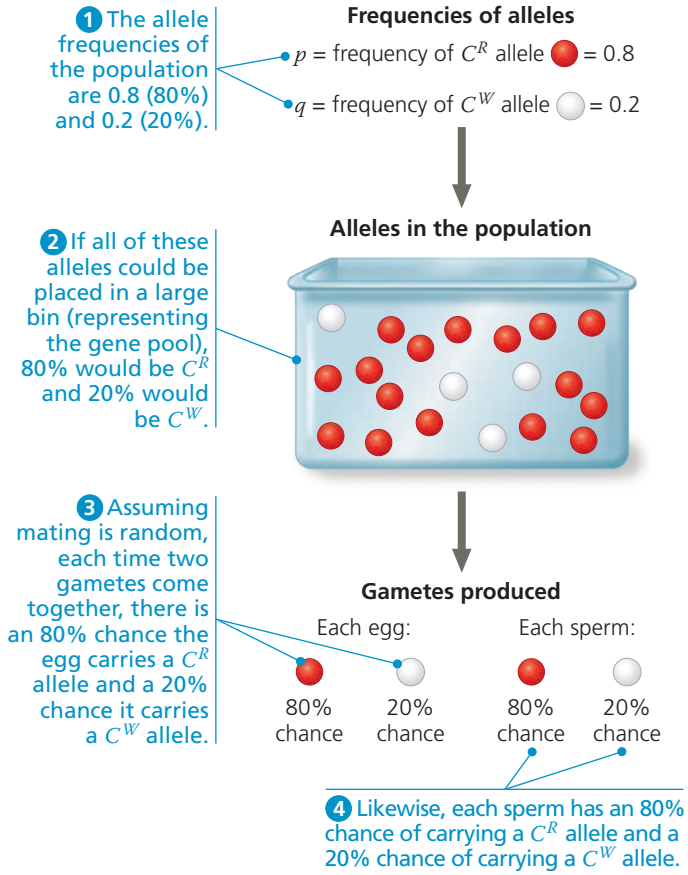
Hardy-Weinberg Equilibrium

In a population that is not evolving, allele and genotype frequencies will remain constant from generation to generation, provided that only Mendelian segregation and recombination of alleles are at work. Such a population is said to be in **Hardy-Weinberg equilibrium**, named for the British mathematician and German physician, respectively, who independently developed this idea in 1908.

To determine whether a population is in Hardy-Weinberg equilibrium, it is helpful to think about genetic crosses in a new way. Previously, we used Punnett squares to determine the genotypes of offspring in a genetic cross (see Figure 14.5). Here, instead of considering the possible allele combinations from one cross, we'll consider the combination of alleles in *all* of the crosses in a population.

Imagine that all the alleles for a given locus from all the individuals in a population are placed in a large bin **(Figure 23.7)**. We can think of this bin as holding the population's gene pool for that locus. "Reproduction" occurs by selecting alleles at random from the bin; somewhat similar events occur in nature when fish release sperm and eggs into the water or when pollen (containing plant sperm) is blown about by the wind. By viewing reproduction as a process of randomly selecting and combining alleles from the bin (the gene pool), we are in effect assuming that mating occurs at random—that is, that all male-female matings are equally likely.

Figure 23.7 Selecting alleles at random from a gene pool.



DRAW IT ➤ *Draw a similar bin that contains six white balls instead of four. For the frequency of* C^R *in the bin to remain equal to 0.8, how many red balls should the bin contain?*

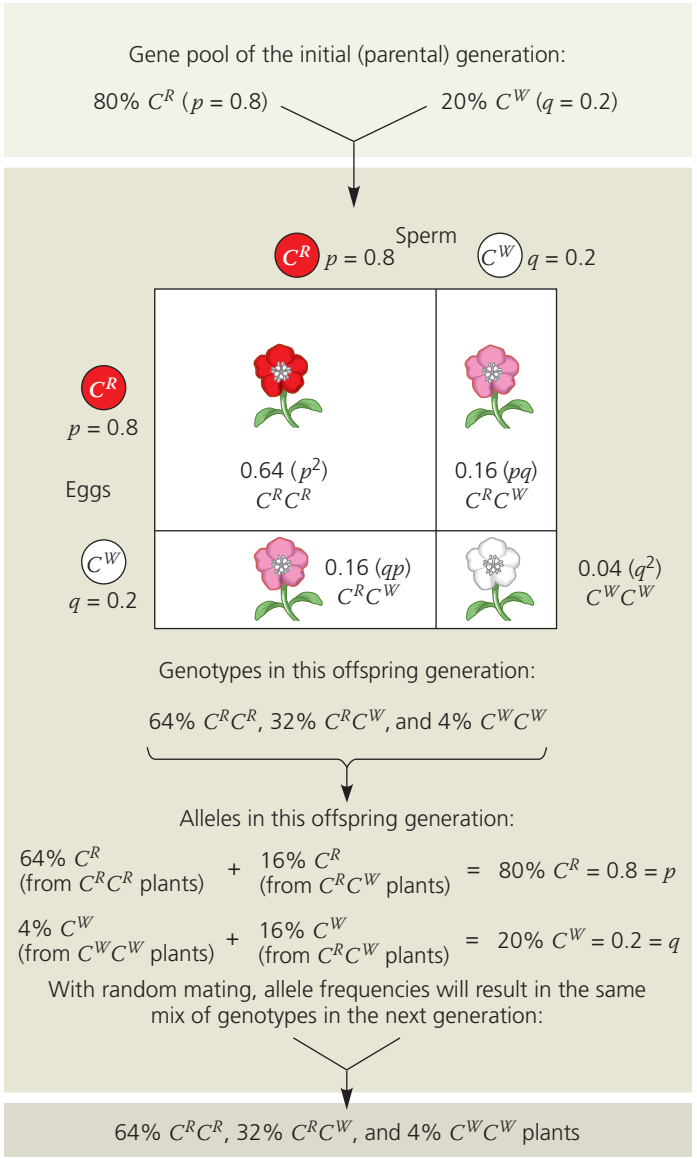
BioFlix® Animation: Allele Frequencies

Let's apply the bin analogy to the hypothetical wildflower population discussed earlier. In that population of 500 flowers, the frequency of the allele for red flowers (C^R) is $p = 0.8$, and the frequency of the allele for white flowers (C^W) is $q = 0.2$. This implies that a bin holding all 1,000 copies of the flower-color gene in the population would contain 800 C^R alleles and 200 C^W alleles. Assuming that gametes are formed by selecting alleles at random from the bin, the probability that an egg or sperm contains a C^R or C^W allele is equal to the frequency of these alleles in the bin. Thus, as shown in Figure 23.7, each egg has an 80% chance of containing a C^R allele and a 20% chance of containing a C^W allele; the same is true for each sperm.

Using the rule of multiplication (see Figure 14.9), we can now calculate the frequencies of the three possible genotypes, assuming random unions of sperm and eggs. The probability that two C^R alleles will come together is $p \times p = p^2 = 0.8 \times 0.8 = 0.64$. Thus, about 64% of the plants in the next generation will have the genotype C^RC^R. The frequency of C^WC^W individuals is expected to be about $q \times q = q^2 = 0.2 \times 0.2 = 0.04$, or 4%. C^RC^W heterozygotes can arise in two different ways. If the sperm provides the C^R allele and the egg provides the C^W allele, the resulting heterozygotes will be $p \times q = 0.8 \times 0.2 = 0.16$, or 16% of the total. If the sperm provides the C^W allele and the egg the C^R allele, the heterozygous offspring will make up $q \times p = 0.2 \times 0.8 = 0.16$, or 16%. The frequency of heterozygotes is thus the sum of these possibilities: $pq + qp = 2pq = 0.16 + 0.16 = 0.32$, or 32%.

As shown in **Figure 23.8**, the genotype frequencies in the next generation must add up to 1 (100%). Thus, the equation for Hardy-Weinberg equilibrium states that at a locus with

Figure 23.8 Hardy-Weinberg equilibrium. In our wildflower population, the gene pool remains constant from one generation to the next. Mendelian processes alone do not alter frequencies of alleles or genotypes.



WHAT IF? ➤ *If the frequency of the* C^R *allele were 0.6, predict the frequencies of the* C^RC^R, C^RC^W, *and* C^WC^W *genotypes.*

Figure Walkthrough

two alleles, the three genotypes will appear in the following proportions:

p^2	+	$2pq$	+	q^2	=	1
Expected frequency of genotype C^RC^R		Expected frequency of genotype C^RC^W		Expected frequency of genotype C^WC^W		

Note that for a locus with two alleles, only three genotypes are possible (in this case, C^RC^R, C^RC^W, and C^WC^W). As a result, the sum of the frequencies of the three genotypes must equal 1 (100%) in *any* population—regardless of whether the population is in Hardy-Weinberg equilibrium. The key point is that a population is in Hardy-Weinberg equilibrium only if the observed genotype frequency of one homozygote is p^2, the observed frequency of the other homozygote is q^2, and the observed frequency of heterozygotes is $2pq$. Finally, as suggested by Figure 23.8, if a population such as our wildflowers is in Hardy-Weinberg equilibrium and its members continue to mate randomly generation after generation, allele and genotype frequencies will remain constant. The system operates somewhat like a deck of cards: No matter how many times the deck is reshuffled to deal out new hands, the deck itself remains the same. Aces do not grow more numerous than jacks. And the repeated shuffling of a population's gene pool over the generations cannot, in itself, change the frequency of one allele relative to another.

Conditions for Hardy-Weinberg Equilibrium

The Hardy-Weinberg approach describes a population that is not evolving. This can occur if a population meets all five of the conditions for Hardy-Weinberg equilibrium listed in **Table 23.1**. But in nature, the allele and genotype frequencies of a population often *do* change over time. Such changes can occur when at least one of the conditions for Hardy-Weinberg equilibrium is not met.

Although departure from the conditions in Table 23.1 is common—resulting in evolutionary change—it is also common for natural populations to be in Hardy-Weinberg equilibrium for specific genes. One way this can happen is if selection alters allele frequencies at some loci but not others. In addition, some populations evolve so slowly that the changes in their allele and genotype frequencies are difficult to distinguish from those predicted for a non-evolving population.

Applying the Hardy-Weinberg Equation

The Hardy-Weinberg equation is often used as an initial test of whether evolution is occurring in a population. The equation also has medical applications, such as estimating the percentage of a population carrying the allele for an inherited disease. For example, consider phenylketonuria (PKU), a metabolic disorder that results from homozygosity for a recessive allele. This disorder occurs in about one out of every 10,000 babies born in the United States. Left untreated, PKU results in mental disability and other problems. (As described in Concept 14.4, newborns are now routinely tested for PKU, and symptoms can be largely avoided with a diet very low in phenylalanine.)

Table 23.1 Conditions for Hardy-Weinberg Equilibrium

Condition	Consequence if Condition Does Not Hold
1. No mutations	The gene pool is modified if mutations occur or if entire genes are deleted or duplicated.
2. Random mating	If individuals mate within a subset of the population, such as near neighbors or close relatives (inbreeding), random mixing of gametes does not occur and genotype frequencies change.
3. No natural selection	Allele frequencies change when individuals with different genotypes show consistent differences in their survival or reproductive success.
4. Extremely large population size	In small populations, allele frequencies fluctuate by chance over time (a process called genetic drift).
5. No gene flow	By moving alleles into or out of populations, gene flow can alter allele frequencies.

Animation: Causes of Evolutionary Change

To apply the Hardy-Weinberg equation, we must assume that no new PKU mutations are being introduced into the population (condition 1) and that people neither choose their mates on the basis of whether or not they carry this gene nor generally mate with close relatives (condition 2). We must also ignore any effects of differential survival and reproductive success among PKU genotypes (condition 3) and assume that there are no effects of genetic drift (condition 4) or of gene flow from other populations into the United States (condition 5). These assumptions are reasonable: The mutation rate for the PKU gene is low, inbreeding and other forms of nonrandom mating are not common in the United States, selection occurs only against the rare homozygotes (and then only if dietary restrictions are not followed), the U.S. population is very large, and populations outside the country have PKU allele frequencies similar to those seen in the United States.

If all these assumptions hold, then the frequency of individuals in the population born with PKU will correspond to q^2 in the Hardy-Weinberg equation (q^2 = frequency of homozygotes). Because the allele is recessive, we must estimate the number of heterozygotes rather than counting them directly as we did with the pink flowers. Recall that there is one PKU occurrence per 10,000 births, which indicates that $q^2 = 0.0001$. Thus, the frequency (q) of the recessive allele for PKU is

$$q = \sqrt{0.0001} = 0.01$$

and the frequency of the dominant allele is

$$p = 1 - q = 1 - 0.01 = 0.99$$

The frequency of carriers, heterozygous people who do not have PKU but may pass the PKU allele to offspring, is

$$2pq = 2 \times 0.99 \times 0.01 = 0.0198$$
(approximately 2% of the U.S. population)

Remember, the assumption of Hardy-Weinberg equilibrium yields an approximation; the real number of carriers may differ. Still, our calculations suggest that harmful recessive alleles at this and other loci can be concealed in a population because they are carried by healthy heterozygotes. The **Scientific Skills Exercise** provides another opportunity for you to apply the Hardy-Weinberg equation to allele data.

CONCEPT CHECK 23.2

1. A population has 700 individuals, 85 of genotype AA, 320 of genotype Aa, and 295 of genotype aa. What are the frequencies of alleles A and a?
2. The frequency of allele a is 0.45 for a population in Hardy-Weinberg equilibrium. What are the expected frequencies of genotypes AA, Aa, and aa?
3. **WHAT IF?** ➤ A locus that affects susceptibility to a degenerative brain disease has two alleles, V and v. In a population, 16 people have genotype VV, 92 have genotype Vv, and 12 have genotype vv. Is this population evolving? Explain.

For suggested answers, see Appendix A.

CONCEPT 23.3

Natural selection, genetic drift, and gene flow can alter allele frequencies in a population

Note again the five conditions required for a population to be in Hardy-Weinberg equilibrium (see Table 23.1). A deviation from any of these conditions is a potential cause of evolution. New mutations (violation of condition 1) can alter allele frequencies, but because mutations are rare, the change from one generation to the next is likely to be very small. Nonrandom mating (violation of condition 2) can affect the frequencies of homozygous and heterozygous genotypes but by itself has no effect on allele frequencies in the gene pool. (Allele frequencies can change if individuals with certain inherited traits are more likely than other individuals to obtain mates. However, such a situation not only causes a deviation from random mating, but also violates condition 3, no natural selection.)

For the rest of this section we will focus on the three mechanisms that alter allele frequencies directly and cause most evolutionary change: natural selection, genetic drift, and gene flow (violations of conditions 3–5).

SCIENTIFIC SKILLS EXERCISE

Using the Hardy-Weinberg Equation to Interpret Data and Make Predictions

Is Evolution Occurring in a Soybean Population? One way to test whether evolution is occurring in a population is to compare the observed genotype frequencies at a locus with those expected for a non-evolving population based on the Hardy-Weinberg equation. In this exercise, you'll test whether a soybean population is evolving at a locus with two alleles, C^G and C^Y, that affect chlorophyll production and hence leaf color.

How the Experiment Was Done Students planted soybean seeds and then counted the number of seedlings of each genotype at day 7 and again at day 21. Seedlings of each genotype could be distinguished visually because the C^G and C^Y alleles show incomplete dominance: C^GC^G seedlings have green leaves, C^GC^Y seedlings have green-yellow leaves, and C^YC^Y seedlings have yellow leaves.

Data from the Experiment

Time (days)	Number of Seedlings			
	Green (C^GC^G)	Green-yellow (C^GC^Y)	Yellow (C^YC^Y)	Total
7	49	111	56	216
21	47	106	20	173

INTERPRET THE DATA

1. Use the observed genotype frequencies from the day 7 data to calculate the frequencies of the C^G allele (p) and the C^Y allele (q).
2. Next, use the Hardy-Weinberg equation ($p^2 + 2pq + q^2 = 1$) to calculate the day 7 expected frequencies of genotypes C^GC^G, C^GC^Y, and C^YC^Y for a population in Hardy-Weinberg equilibrium.
3. Calculate the observed frequencies of genotypes C^GC^G, C^GC^Y, and C^YC^Y at day 7. Compare these frequencies to the expected frequencies calculated in question 2. Is the seedling population in Hardy-Weinberg equilibrium at day 7, or is evolution occurring? Explain your reasoning and identify which genotypes, if any, appear to be selected for or against.
4. Calculate the observed frequencies of genotypes C^GC^G, C^GC^Y, and C^YC^Y at day 21. Compare these frequencies to the expected frequencies calculated in question 2 and to the observed frequencies at day 7. Is the seedling population in Hardy-Weinberg equilibrium at day 21, or is evolution occurring? Explain your reasoning and identify which genotypes, if any, appear to be selected for or against.
5. Homozygous C^YC^Y individuals cannot produce chlorophyll. The ability to photosynthesize becomes more critical as seedlings age and begin to exhaust the supply of food that was stored in the seed from which they emerged. Develop a hypothesis that explains the data for days 7 and 21. Based on this hypothesis, predict how the frequencies of the C^G and C^Y alleles will change beyond day 21.

MB **Instructors**: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Natural Selection

The concept of natural selection is based on differential success in survival and reproduction: Individuals in a population exhibit variations in their heritable traits, and those with traits that are better suited to their environment tend to produce more offspring than those with traits that are not as well suited.

In genetic terms, selection results in alleles being passed to the next generation in proportions that differ from those in the present generation. For example, the fruit fly *D. melanogaster* has an allele that confers resistance to several insecticides, including DDT. This allele has a frequency of 0% in laboratory strains of *D. melanogaster* established from flies collected in the wild in the early 1930s, prior to DDT use. However, in strains established from flies collected after 1960 (following 20 or more years of DDT use), the allele frequency is 37%. We can infer that this allele either arose by mutation between 1930 and 1960 or was present in 1930, but very rare. In any case, the rise in frequency of this allele most likely occurred because DDT is a powerful poison that is a strong selective force in exposed fly populations.

As the *D. melanogaster* example suggests, an allele that confers resistance to an insecticide will increase in frequency in a population exposed to that insecticide. Such changes are not coincidental. By consistently favoring some alleles over others, natural selection can cause **adaptive evolution**, a process in which traits that enhance survival or reproduction tend to increase in frequency over time. We'll explore this process in more detail later in this chapter.

Genetic Drift

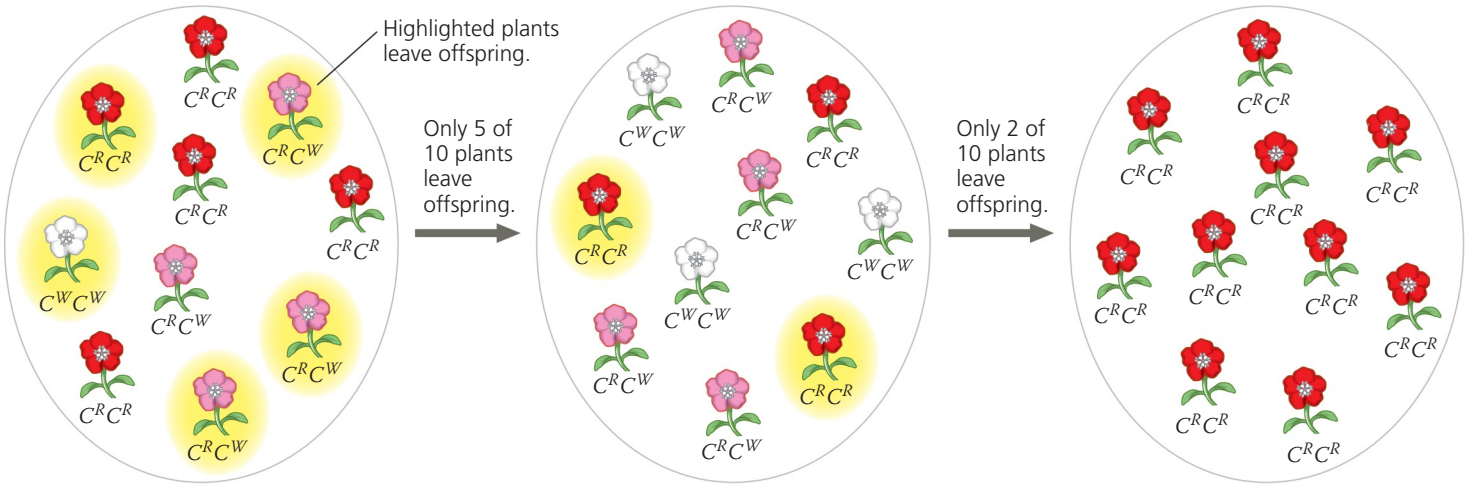
If you flip a coin 1,000 times, a result of 700 heads and 300 tails might make you suspicious about that coin. But if you flip a coin only 10 times, an outcome of 7 heads and 3 tails would not be surprising. The smaller the number of coin flips, the more likely it is that chance alone will cause a deviation from the predicted result. (In this case, the prediction is an equal number of heads and tails.) Chance events can also cause allele frequencies to fluctuate unpredictably from one generation to the next, especially in small populations—a process called **genetic drift**.

Figure 23.9 models how genetic drift might affect a small population of our wildflowers. In this example, drift leads to the loss of an allele from the gene pool, but it is a matter of chance that the C^W allele is lost and not the C^R allele. Such unpredictable changes in allele frequencies can be caused by chance events associated with survival and reproduction. Perhaps a large animal such as a moose stepped on the three C^WC^W individuals in generation 2, killing them and increasing the chance that only the C^R allele would be passed to the next generation. Allele frequencies can also be affected by chance events that occur during fertilization. For example, suppose two individuals of genotype C^RC^W had a small number of offspring. By chance alone, every egg and sperm pair that generated offspring could happen to have carried the C^R allele and not the C^W allele.

Certain circumstances can result in genetic drift having a significant impact on a population. Two examples are the founder effect and the bottleneck effect.

Figure 23.9 Genetic drift. This small wildflower population has a stable size of ten plants. Suppose that by chance only five plants of generation 1 (those highlighted in yellow) produce fertile offspring. (This could occur, for example, if only those plants happened to grow in a location that provided enough nutrients to support the production of offspring.) Again by chance, only two plants of generation 2 leave fertile offspring.

VISUAL SKILLS ➤ *Based on this diagram, summarize how the frequency of the C^W allele changes over time.*



The Founder Effect

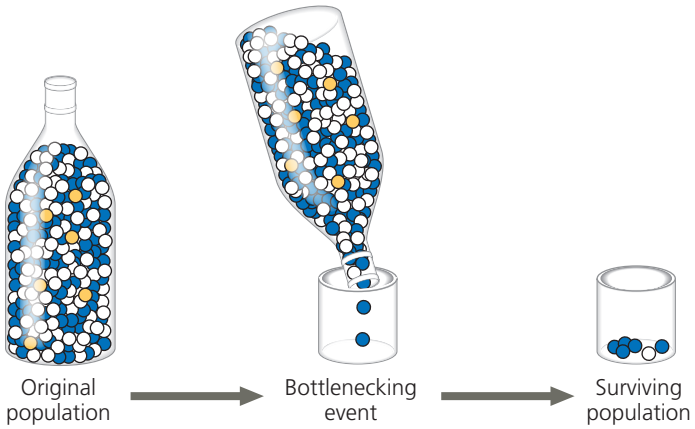
When a few individuals become isolated from a larger population, this smaller group may establish a new population whose gene pool differs from the source population; this is called the **founder effect**. The founder effect might occur, for example, when a few members of a population are blown by a storm to a new island. Genetic drift, in which chance events alter allele frequencies, can occur in such a case because the storm indiscriminately transports some individuals (and their alleles), but not others, from the source population.

The founder effect probably accounts for the relatively high frequency of certain inherited disorders among isolated human populations. For example, in 1814, 15 British colonists founded a settlement on Tristan da Cunha, a group of small islands in the Atlantic Ocean midway between Africa and South America. Apparently, one of the colonists carried a recessive allele for retinitis pigmentosa, a progressive form of blindness that afflicts homozygous individuals. Of the founding colonists' 240 descendants on the island in the late 1960s, four had retinitis pigmentosa. The frequency of the allele that causes this disease is ten times higher on Tristan da Cunha than in the populations from which the founders came.

The Bottleneck Effect

A sudden change in the environment, such as a fire or flood, may drastically reduce the size of a population. A severe drop in population size can cause the **bottleneck effect**, so named because the population has passed through a "bottleneck" that reduces its size **(Figure 23.10)**. By chance alone, certain alleles may be overrepresented among the survivors, others may be underrepresented, and some may be absent altogether. Ongoing genetic drift is likely to have substantial effects on the gene pool until the population becomes large enough that chance events have less impact. But even if a population that has passed through a bottleneck ultimately recovers in size, it may have low levels of genetic variation for a long period of time—a legacy of the genetic drift that occurred when the population was small.

Figure 23.10 The bottleneck effect. Shaking just a few marbles through the narrow neck of a bottle is analogous to a drastic reduction in the size of a population. By chance, blue marbles are overrepresented in the surviving population and gold marbles are absent.

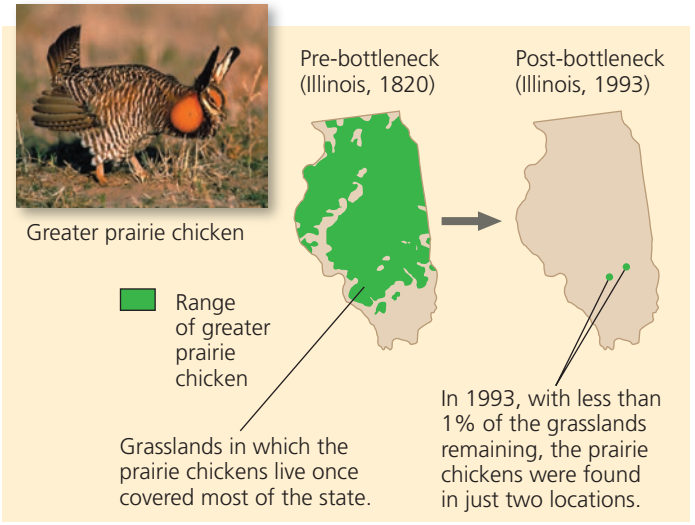


Human actions sometimes create severe bottlenecks for other species, as the following example shows.

Case Study: Impact of Genetic Drift on the Greater Prairie Chicken

Millions of greater prairie chickens (*Tympanuchus cupido*) once lived on the prairies of Illinois. As these prairies were converted to farmland and other uses during the 19th and 20th centuries, the number of greater prairie chickens plummeted **(Figure 23.11a)**. By 1993 fewer than 50 birds remained. These few surviving birds had low levels of genetic variation, and less than 50% of their eggs hatched, compared with much higher hatching rates of the larger populations in Kansas and Nebraska **(Figure 23.11b)**.

Figure 23.11 Genetic drift and loss of genetic variation.



(a) The Illinois population of greater prairie chickens dropped from millions of birds in the 1800s to fewer than 50 birds in 1993.

Location	Population size	Number of alleles per locus	Percentage of eggs hatched
Illinois			
1930–1960s	1,000–25,000	5.2	93
1993	<50	3.7	<50
Kansas, 1998 (no bottleneck)	750,000	5.8	99
Nebraska, 1998 (no bottleneck)	75,000–200,000	5.8	96

(b) In the small Illinois population, genetic drift led to decreases in the number of alleles per locus and the percentage of eggs hatched.

These data suggest that genetic drift during the bottleneck may have led to a loss of genetic variation and an increase in the frequency of harmful alleles. To investigate this hypothesis, researchers extracted DNA from 15 museum specimens of Illinois greater prairie chickens. Of the 15 birds, 10 had been collected in the 1930s, when there were 25,000 greater prairie chickens in Illinois, and 5 had been collected in the 1960s, when there were 1,000 greater prairie chickens in Illinois. By studying the DNA of these specimens, the researchers were able to obtain a minimum, baseline estimate of how much genetic variation was present in the Illinois population *before* the population shrank to extremely low numbers. This baseline estimate is a key piece of information that is not usually available in cases of population bottlenecks.

The researchers surveyed six loci and found that the 1993 population had fewer alleles per locus than the pre-bottleneck Illinois or the current Kansas and Nebraska populations (see Figure 23.11b). Thus, as predicted, drift had reduced the genetic variation of the small 1993 population. Drift may also have increased the frequency of harmful alleles, leading to the low egg-hatching rate. To counteract these negative effects, 271 birds from neighboring states were added to the Illinois population over four years. This strategy succeeded: New alleles entered the population, and the egg-hatching rate improved to over 90%. Overall, studies on the Illinois greater prairie chicken illustrate the powerful effects of genetic drift in small populations and provide hope that in at least some populations, these effects can be reversed.

Effects of Genetic Drift: *A Summary*

The examples we've described highlight four key points:

1. **Genetic drift is significant in small populations.** Chance events can cause an allele to be disproportionately over- or underrepresented in the next generation. Although chance events occur in populations of all sizes, they tend to alter allele frequencies substantially only in small populations.
2. **Genetic drift can cause allele frequencies to change at random.** Because of genetic drift, an allele may increase in frequency one year, then decrease the next; the change from year to year is not predictable. Thus, unlike natural selection, which in a given environment consistently favors some alleles over others, genetic drift causes allele frequencies to change at random over time.
3. **Genetic drift can lead to a loss of genetic variation within populations.** By causing allele frequencies to fluctuate randomly over time, genetic drift can eliminate alleles from a population. Because evolution depends on genetic variation, such losses can influence how effectively a population can adapt to a change in the environment.
4. **Genetic drift can cause harmful alleles to become fixed.** Alleles that are neither harmful nor beneficial can be lost or become fixed (reach a frequency of 100%) by chance through genetic drift. In very small populations, genetic drift can also cause alleles that are slightly harmful to become fixed. When this occurs, the population's survival can be threatened (as in greater prairie chickens).

Gene Flow

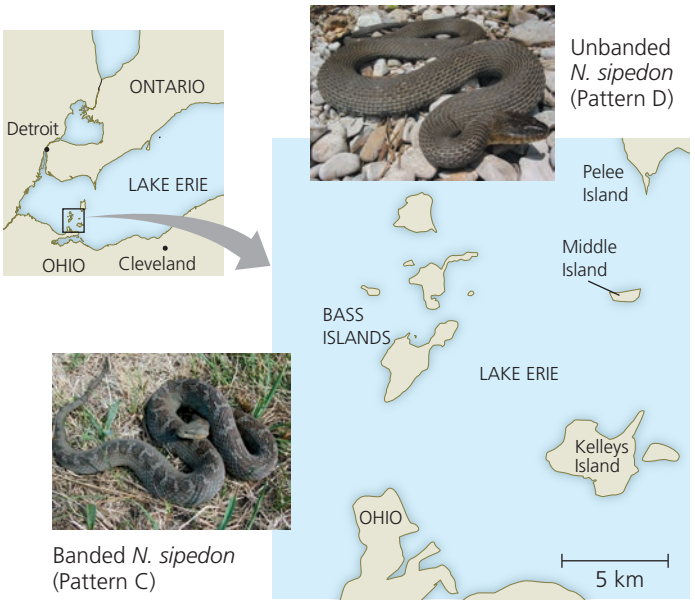
Natural selection and genetic drift are not the only phenomena affecting allele frequencies. Allele frequencies can also change by **gene flow**, the transfer of alleles into or out of a population due to the movement of fertile individuals or their gametes. For example, suppose that near our original hypothetical wildflower population there is another population consisting primarily of white-flowered individuals (C^WC^W). Insects carrying pollen from these plants may fly to and pollinate plants in our original population. The introduced C^W alleles would modify our original population's allele frequencies in the next generation. Because alleles are transferred between populations, gene flow tends to reduce the genetic differences between populations. In fact, if it is extensive enough, gene flow can result in two populations combining into a single population with a common gene pool.

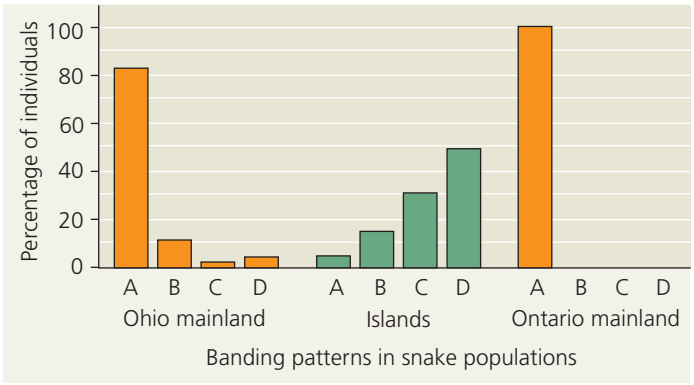
Alleles transferred by gene flow can also affect how well populations are adapted to local environmental conditions. For instance, mainland and island populations of the Lake Erie water snake (*Nerodia sipedon*) differ in their color patterns: Nearly all snakes from the Ohio or Ontario mainlands are strongly banded, whereas the majority of snakes from islands are unbanded or intermediate **(Figure 23.12)**. Banding coloration is an inherited trait, determined by a few loci (with alleles that encode bands being dominant to alleles that encode the absence of bands). On islands, water snakes live along rocky shorelines, while on the mainland, they live in marshes. Snakes without bands are more well camouflaged in island habitats than are snakes with bands. Hence, on islands, snakes without bands survive at higher rates than do snakes with bands.

These data indicate that snakes without bands are favored by natural selection in island populations. Thus, we might expect that *all* snakes on islands would lack bands. Why is this not the case? The answer lies in gene flow from the mainland. In any given year, 3 to10 snakes from the mainland swim to the islands and join the populations there. As a result, each year such migrants transfer alleles for banded coloration from the mainland (where nearly all snakes have bands) to the islands. This ongoing gene flow has prevented selection from removing all of the alleles for banded coloration from island populations—thereby preventing island populations from adapting fully to local conditions.

Gene flow can also transfer alleles that improve the ability of populations to adapt to local conditions. For example, gene flow has resulted in the worldwide spread of several insecticide resistance alleles in the mosquito *Culex pipiens*,

Figure 23.12 Gene flow and local adaptation in the Lake Erie water snake (*Nerodia sipedon*). Researchers assigned letters to variations in coloration in *N. sipedon* populations. Color pattern A is strong banding, patterns B and C are intermediate banding, and pattern D is no banding. Banding is advantageous for camouflage in mainland environments, whereas having no bands is advantageous in island environments. However, gene flow from the mainland causes banding to persist in island populations.





WHAT IF? ➤ *Suppose a severe weather event caused island populations to decrease in size but did not affect the size of mainland populations. Predict how gene flow from the mainland would affect color patterns in island populations. Explain.*

a vector of West Nile virus and other diseases. Each of these alleles has a unique genetic signature that allowed researchers to document that it arose by mutation in only one or a few geographic locations. In their population of origin, these alleles increased because they provided insecticide resistance. These alleles were then transferred to new populations, where again, their frequencies increased as a result of natural selection.

Finally, gene flow has become an increasingly important agent of evolutionary change in human populations. Humans today move much more freely about the world than in the past. As a result, mating is more common between members of populations that previously had very little contact, leading to an exchange of alleles and fewer genetic differences between those populations.

BioFlix® Animation: Natural Selection, Genetic Drift, and Gene Flow

CONCEPT CHECK 23.3

1. In what sense is natural selection more "predictable" than genetic drift?
2. Distinguish genetic drift from gene flow in terms of (a) how they occur and (b) their implications for future genetic variation in a population.
3. **WHAT IF?** ➤ Suppose two plant populations exchange pollen and seeds. In one population, individuals of genotype *AA* are most common (9,000 *AA*, 900 *Aa*, 100 *aa*), while the opposite is true in the other population (100 *AA*, 900 *Aa*, 9,000 *aa*). If neither allele has a selective advantage, what will happen over time to the allele and genotype frequencies of these populations?

For suggested answers, see Appendix A.

CONCEPT 23.4

Natural selection is the only mechanism that consistently causes adaptive evolution

Evolution by natural selection is a blend of chance and "sorting": chance in the creation of new genetic variations (as in mutation) and sorting as natural selection favors some alleles over others. Because of this favoring process, the outcome of natural selection is *not* random. Instead, natural selection consistently increases the frequencies of alleles that provide reproductive advantage, thus leading to adaptive evolution.

Natural Selection: *A Closer Look*

To see how natural selection can cause adaptive evolution, we'll begin with the concept of relative fitness and the different ways that selection acts on an organism's phenotype.

Relative Fitness

The phrases "struggle for existence" and "survival of the fittest" are commonly used to describe natural selection, but these expressions are misleading if always taken to mean direct competitive contests among individuals. There *are* animal species in which individuals, usually the males, lock horns or otherwise do combat to determine mating privilege. But reproductive success is generally more subtle and depends on many factors besides outright battle. For example, a barnacle that is more efficient at collecting food than its neighbors may have greater stores of energy and hence be able to produce a larger number of eggs. A moth may have more offspring than other moths in the same population because its body colors more effectively conceal

it from predators, improving its chance of surviving long enough to produce more offspring. These examples illustrate how in a given environment, certain traits can lead to greater **relative fitness**: the contribution an individual makes to the gene pool of the next generation *relative to* the contributions of other individuals.

Although we often refer to the relative fitness of a genotype, remember that the entity that is subjected to natural selection is the whole organism, not the underlying genotype. Thus, selection acts more directly on the phenotype than on the genotype; it acts on the genotype indirectly, via how the genotype affects the phenotype.

Directional, Disruptive, and Stabilizing Selection

Natural selection can alter the frequency distribution of heritable traits in three ways, depending on which phenotypes in a population are favored: through directional selection, disruptive selection, and stabilizing selection.

Directional selection occurs when conditions favor individuals exhibiting one extreme of a phenotypic range, thereby shifting a population's frequency curve for the phenotypic character in one direction or the other **(Figure 23.13a)**. Directional selection is common when a population's environment changes or when members of a population migrate to a new (and different) habitat. For instance, an increase in the relative abundance of large seeds over small seeds led to an increase in beak depth in a population of Galápagos finches (see Figure 23.2).

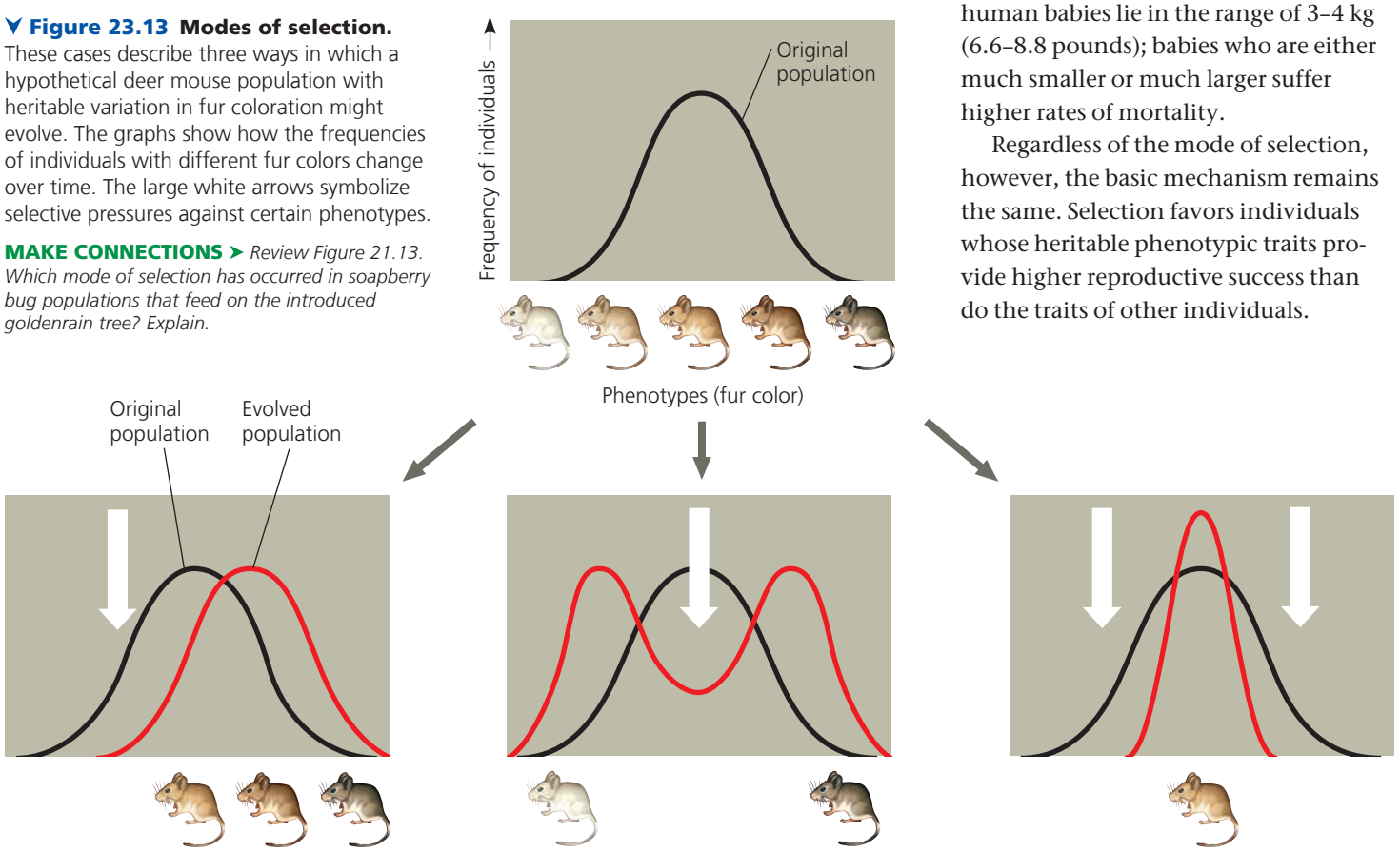
Disruptive selection (Figure 23.13b) occurs when conditions favor individuals at both extremes of a phenotypic range over individuals with intermediate phenotypes. One example is a population of black-bellied seedcracker finches in Cameroon whose members display two distinctly different beak sizes. Small-billed birds feed mainly on soft seeds, whereas large-billed birds specialize in cracking hard seeds. It appears that birds with intermediate-sized bills are relatively inefficient at cracking both types of seeds and thus have lower relative fitness.

Stabilizing selection (Figure 23.13c) acts against both extreme phenotypes and favors intermediate variants. This mode of selection reduces variation and tends to maintain the status quo for a particular phenotypic character. For example, the birth weights of most human babies lie in the range of 3–4 kg (6.6–8.8 pounds); babies who are either much smaller or much larger suffer higher rates of mortality.

Regardless of the mode of selection, however, the basic mechanism remains the same. Selection favors individuals whose heritable phenotypic traits provide higher reproductive success than do the traits of other individuals.

Figure 23.13 Modes of selection. These cases describe three ways in which a hypothetical deer mouse population with heritable variation in fur coloration might evolve. The graphs show how the frequencies of individuals with different fur colors change over time. The large white arrows symbolize selective pressures against certain phenotypes.

MAKE CONNECTIONS ➤ *Review Figure 21.13. Which mode of selection has occurred in soapberry bug populations that feed on the introduced goldenrain tree? Explain.*



(a) Directional selection shifts the overall makeup of the population by favoring variants that are at one extreme of the distribution. In this case, lighter mice are selected against because they live among dark rocks, making it harder for them to hide from predators.

(b) Disruptive selection favors variants at both ends of the distribution. These mice have colonized a patchy habitat made up of light and dark rocks, with the result that mice of an intermediate color are selected against.

(c) Stabilizing selection removes extreme variants from the population and preserves intermediate types. If the environment consists of rocks of an intermediate color, both light and dark mice will be selected against.

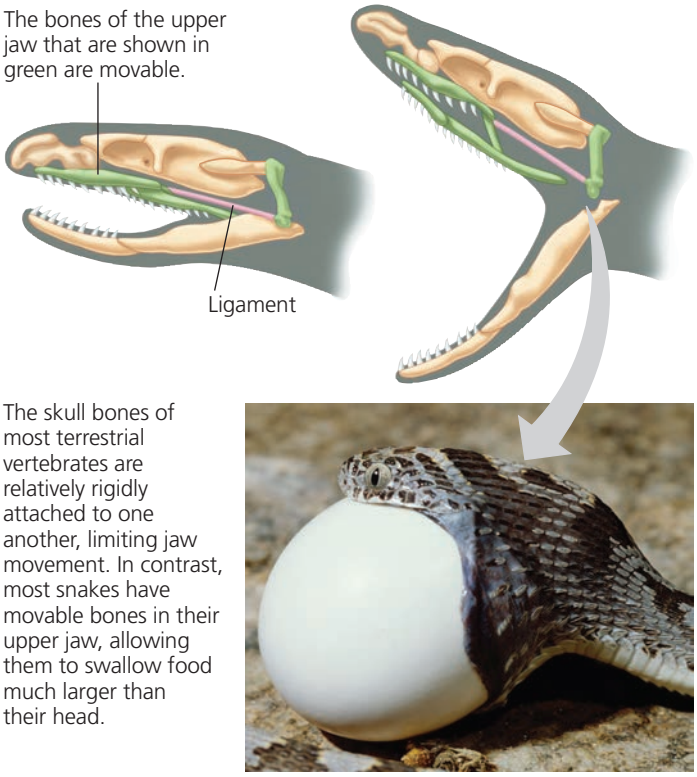
The Key Role of Natural Selection in Adaptive Evolution

The adaptations of organisms include many striking examples. Certain octopuses, for example, have the ability to change color rapidly, enabling them to blend into different backgrounds. Another example is the remarkable jaws of snakes **(Figure 23.14)**, which allow them to swallow prey much larger than their own head (a feat analogous to a person swallowing a whole watermelon). Other adaptations, such as a version of an enzyme that shows improved function in cold environments, may be less visually dramatic but just as important for survival and reproduction.

Such adaptations can arise gradually over time as natural selection increases the frequencies of alleles that enhance survival or reproduction. As the proportion of individuals that have favorable traits increases, the degree to which a species is well suited for life in its environment improves; that is, adaptive evolution occurs. However, the physical and biological components of an organism's environment may change over time. As a result, what constitutes a "good match" between an organism and its environment can be a moving target, making adaptive evolution a continuous, dynamic process. Environmental conditions can also differ from place to place, causing different alleles to be favored in different locations. When this occurs, natural selection can cause the populations of a species to differ genetically from one another.

And what about genetic drift and gene flow? Both can, in fact, increase the frequencies of alleles that enhance survival or reproduction, but neither does so consistently. Genetic drift can cause the frequency of a slightly beneficial allele to increase, but it also can cause the frequency of such an allele to decrease. Similarly, gene flow may introduce alleles that are advantageous or ones that are disadvantageous. Natural selection is the only evolutionary mechanism that consistently leads to adaptive evolution.

Figure 23.14 Movable jaw bones in snakes.



HHMI Video: Got Lactase? The Co-evolution of Genes and Culture

Sexual Selection

Charles Darwin was the first to explore the implications of **sexual selection**, a process in which individuals with certain inherited characteristics are more likely than other individuals of the same sex to obtain mates. Sexual selection can result in **sexual dimorphism**, a difference in secondary sexual characteristics between males and females of the same species **(Figure 23.15)**. These distinctions include differences in size, color, ornamentation, and behavior.

How does sexual selection operate? There are several ways. In **intrasexual selection**, meaning selection within the same sex, individuals of one sex compete directly for mates of the opposite sex. In many species, intrasexual selection occurs among males. For example, a single male may patrol a group of females and prevent other males from mating with them. The patrolling male may defend his status by defeating smaller, weaker, or less fierce males in combat. More often, this male is the psychological victor in ritualized displays that discourage would-be competitors but do not risk injury that would reduce his own fitness (see Figure 52.16). Intrasexual selection also occurs among females in a variety of species, including ring-tailed lemurs and broadnosed pipefish.

In **intersexual selection**, also called *mate choice*, individuals of one sex (usually the females) are choosy in selecting their mates from the other sex. In many cases, the female's choice depends on the showiness of the male's appearance

Figure 23.15 Sexual dimorphism and sexual selection. A peacock (left) and a peahen (right) show extreme sexual dimorphism. There is intrasexual selection between competing males, followed by intersexual selection when the females choose among the showiest males.

or behavior (see Figure 23.15). What intrigued Darwin about mate choice is that male showiness may not seem adaptive in any other way and may in fact pose some risk. For example, bright plumage may make male birds more visible to predators. But if such characteristics help a male gain a mate, and if this benefit outweighs the risk from predation, then both the bright plumage and the female preference for it will be reinforced because they enhance overall reproductive success.

How do female preferences for certain male characteristics evolve in the first place? One hypothesis is that females prefer male traits that are correlated with "good genes." If the trait preferred by females is indicative of a male's overall genetic quality, both the male trait and female preference for it should increase in frequency. **Figure 23.16** describes one experiment testing this hypothesis in gray tree frogs.

Other researchers have shown that in several bird species, the traits preferred by females are related to overall male health. Here, too, female preference appears to be based on traits that reflect "good genes," in this case, alleles indicative of a robust immune system.

Balancing Selection

As we've seen, genetic variation is often found at loci affected by selection. What prevents natural selection from reducing the variation at those loci by culling all unfavorable alleles? As mentioned earlier, in diploid organisms, many unfavorable recessive alleles persist because they are hidden from selection when in heterozygous individuals. In addition, selection itself may preserve variation at some loci, thus maintaining two or more phenotypic forms in a population. Known as **balancing selection**, this type of selection includes frequency-dependent selection and heterozygote advantage.

Frequency-Dependent Selection

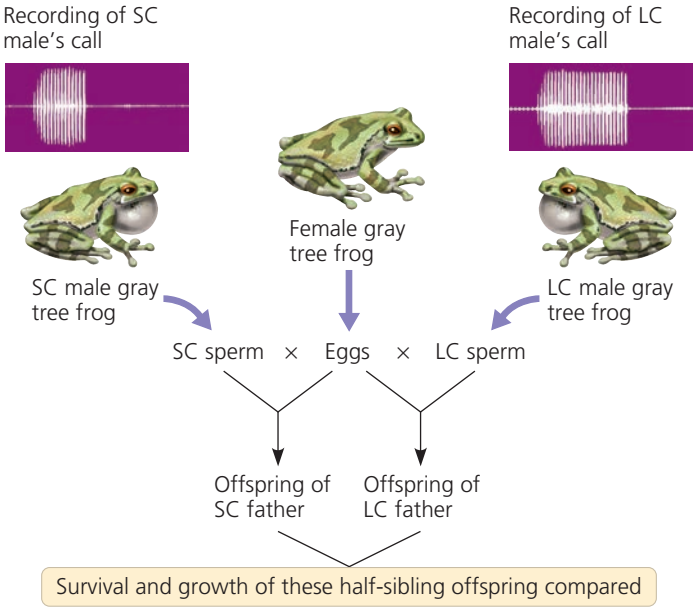
In **frequency-dependent selection**, the fitness of a phenotype depends on how common it is in the population. Consider the scale-eating fish (*Perissodus microlepis*) of Lake Tanganyika, in Africa. These fish attack other fish from behind, darting in to remove a few scales from the flank of their prey. Of interest here is a peculiar feature of the scale-eating fish: Some are "left-mouthed" and some are "right-mouthed." This trait is determined by two alleles and simple Mendelian inheritance. Hence, all individuals in a population are either left-mouthed or right-mouthed, and the frequencies of these two phenotypes must add up to 100%.

Because their mouth twists to the left, left-mouthed fish always attack their prey's right flank **(Figure 23.17)**. (To see why, twist your lower jaw and lips to the left and imagine trying to take a bite from the left side of a fish, approaching it from behind.) Similarly, right-mouthed fish always attack from the left. Prey species guard against attack from whatever phenotype of scale-eating fish is most common in the lake.

Figure 23.16

Inquiry Do females select mates based on traits indicative of "good genes"?

Experiment Female gray tree frogs (*Hyla versicolor*) prefer to mate with males that give long mating calls. Allison Welch and colleagues, at the University of Missouri, tested whether the genetic makeup of long-calling (LC) males is superior to that of short-calling (SC) males. The researchers fertilized half the eggs of each female with sperm from an LC male and fertilized the remaining eggs with sperm from an SC male. In two separate experiments (one in 1995, the other in 1996), the resulting half-sibling offspring were raised in a common environment and their survival and growth were monitored.



Results

Offspring Performance	1995	1996
Larval survival	LC better	NSD
Larval growth	NSD	LC better
Time to metamorphosis	LC better (shorter)	LC better (shorter)

NSD = no significant difference; LC better = offspring of LC males superior to offspring of SC males.

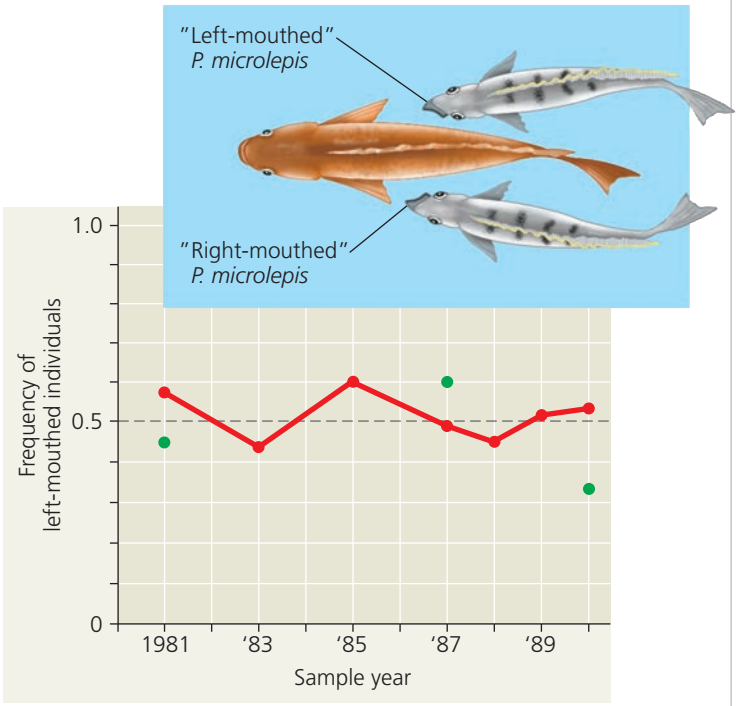
Conclusion Because offspring fathered by an LC male outperformed their half-siblings fathered by an SC male, the team concluded that the duration of a male's mating call is indicative of the male's overall genetic quality. This result supports the hypothesis that female mate choice can be based on a trait that indicates whether the male has "good genes."

Data from A. M. Welch et al., Call duration as an indicator of genetic quality in male gray tree frogs, *Science* 280:1928–1930 (1998).

INQUIRY IN ACTION Read and analyze the original paper in *Inquiry in Action: Interpreting Scientific Papers.*

WHAT IF? *Why did the researchers split each female frog's eggs into two batches for fertilization by different males? Why didn't they mate each female with a single male frog?*

Figure 23.17 Frequency-dependent selection. In a population of the scale-eating fish *Perissodus microlepis*, the frequency of left-mouthed individuals (red data points) rises and falls in a regular manner. The frequency of left-mouthed individuals among adults that reproduced was also recorded in three sample years (green data points).



INTERPRET THE DATA ➤ *For 1981, 1987, and 1990, compare the frequency of left-mouthed individuals among breeding adults to the frequency of left-mouthed individuals in the entire population. What do the data indicate about when natural selection favors left-mouthed individuals over right-mouthed individuals (or vice versa)? Explain.*

Thus, from year to year, selection favors whichever mouth phenotype is least common. As a result, the frequency of left- and right-mouthed fish oscillates over time, and balancing selection (due to frequency dependence) keeps the frequency of each phenotype close to 50%.

Heterozygote Advantage

If individuals who are heterozygous at a particular locus have greater fitness than do both kinds of homozygotes, they exhibit **heterozygote advantage**. In such a case, natural selection tends to maintain two or more alleles at that locus. Note that heterozygote advantage is defined in terms of *genotype*, not phenotype. Thus, whether heterozygote advantage represents stabilizing or directional selection depends on the relationship between the genotype and the phenotype. For example, if the phenotype of a heterozygote is intermediate to the phenotypes of both homozygotes, heterozygote advantage is a form of stabilizing selection.

An example of heterozygote advantage occurs at the locus in humans that codes for the β polypeptide subunit of hemoglobin, the oxygen-carrying protein of red blood cells. In homozygous individuals, a recessive allele at that locus causes sickle-cell disease. The red blood cells of people with sickle-cell disease become distorted in shape, or *sickled*, under low-oxygen conditions (see Figure 5.19), as occurs in the capillaries. These sickled cells can clump together and block the flow of blood in the capillaries, damaging organs such as the kidney, heart, and brain. Although some red blood cells become sickled in heterozygotes, not enough become sickled to cause sickle-cell disease.

Heterozygotes for the sickle-cell allele are protected against the most severe effects of malaria, a disease caused by a parasite that infects red blood cells (see Figure 28.16). One reason for this partial protection is that the body destroys sickled red blood cells rapidly, killing the parasites they harbor. Malaria is a major killer in some tropical regions. In such regions, selection favors heterozygotes over homozygous dominant individuals, who are more vulnerable to the effects of malaria, and also over homozygous recessive individuals, who develop sickle-cell disease. As described in **Figure 23.18**, these selective pressures have caused the frequency of the sickle-cell allele to reach relatively high levels in areas where the malaria parasite is common.

Why Natural Selection Cannot Fashion Perfect Organisms

Though natural selection leads to adaptation, nature abounds with examples of organisms that are less than ideally suited for their lifestyles. There are several reasons why.

1. **Selection can act only on existing variations.** Natural selection favors only the fittest phenotypes among those currently in the population, which may not be the ideal traits. New advantageous alleles do not arise on demand.
2. **Evolution is limited by historical constraints.** Each species has a legacy of descent with modification from ancestral forms. Evolution does not scrap the ancestral anatomy and build each new complex structure from scratch; rather, evolution co-opts existing structures and adapts them to new situations. We could imagine that if a terrestrial animal were to adapt to an environment in which flight would be advantageous, it might be best just to grow an extra pair of limbs that would serve as wings. However, evolution does not work this way; instead, it operates on the traits an organism already has. Thus, in birds and bats, an existing pair of limbs took on new functions for flight as these organisms evolved from nonflying ancestors.
3. **Adaptations are often compromises.** Each organism must do many different things. A seal spends part of its time on rocks; it could probably walk better if it had legs instead of flippers, but then it would not swim nearly as well. We humans owe much of our versatility and athleticism to our prehensile hands and flexible limbs, but these also make us prone to sprains, torn ligaments, and dislocations: Structural reinforcement has been compromised for agility.

Figure 23.18 **MAKE CONNECTIONS**

The Sickle-Cell Allele

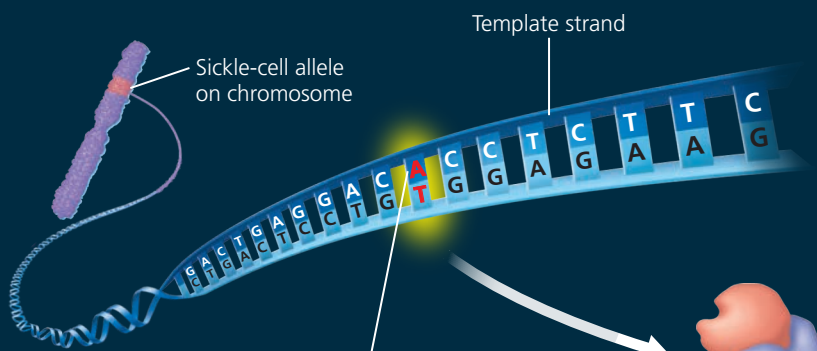
This child has sickle-cell disease, a genetic disorder that strikes individuals that have two copies of the sickle-cell allele. This allele causes an abnormality in the structure and function of hemoglobin, the oxygen-carrying protein in red blood cells. Although sickle-cell disease is lethal if not treated, in some regions the sickle-cell allele can reach frequencies as high as 15–20%. How can such a harmful allele be so common?

Events at the Molecular Level

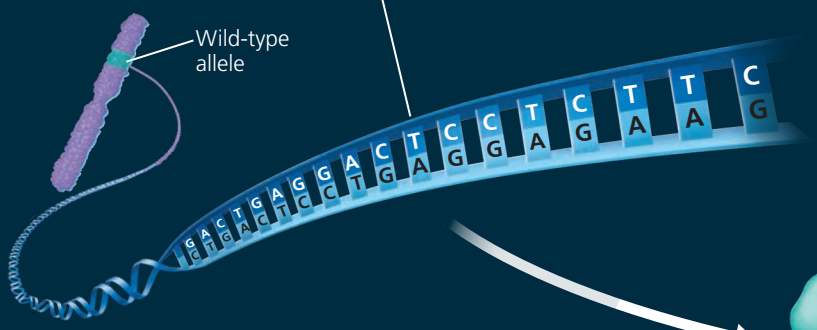
- Due to a point mutation, the sickle-cell allele differs from the wild-type allele by a single nucleotide. (See Figure 17.26.)
- The resulting change in one amino acid leads to hydrophobic interactions between the sickle-cell hemoglobin proteins under low-oxygen conditions.
- As a result, the sickle-cell proteins bind to each other in chains that together form a fiber.

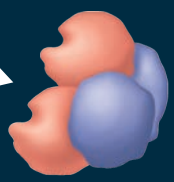
Consequences for Cells

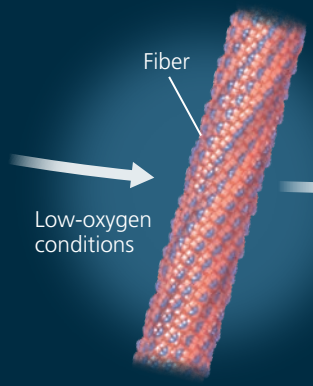
- The abnormal hemoglobin fibers distort the red blood cell into a sickle shape under low-oxygen conditions, such as those found in blood vessels returning to the heart.

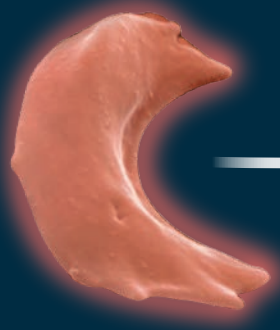


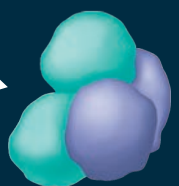
An adenine replaces a thymine in the template strand of the sickle-cell allele, changing one codon in the mRNA produced during transcription. This change causes an amino-acid change in sickle-cell hemoglobin: A valine replaces a glutamic acid at one position. (See Figure 5.19.)

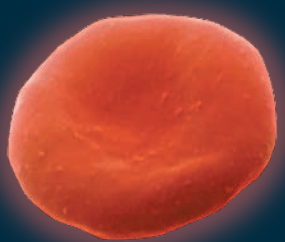




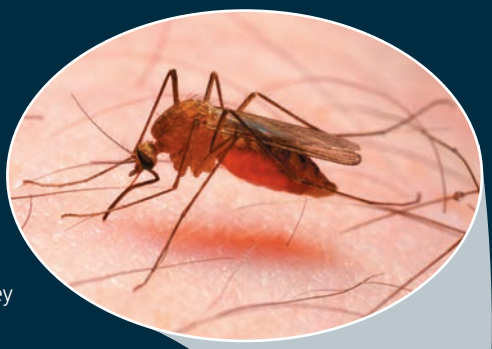








HHMI Video: The Making of the Fittest: Natural Selection in Humans (Sickle-Cell Disease)



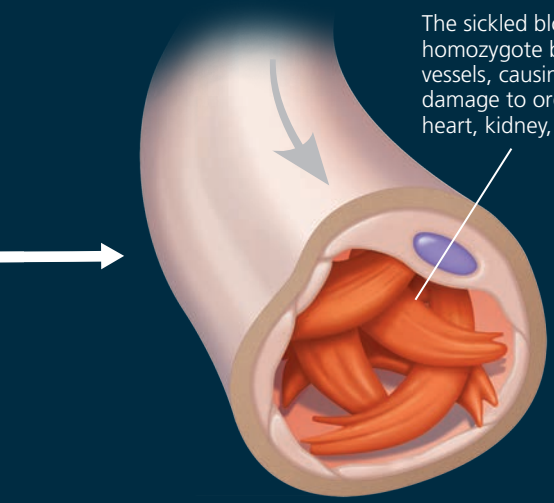
Infected mosquitoes spread malaria when they bite people. (See Figure 28.16.)

Evolution in Populations

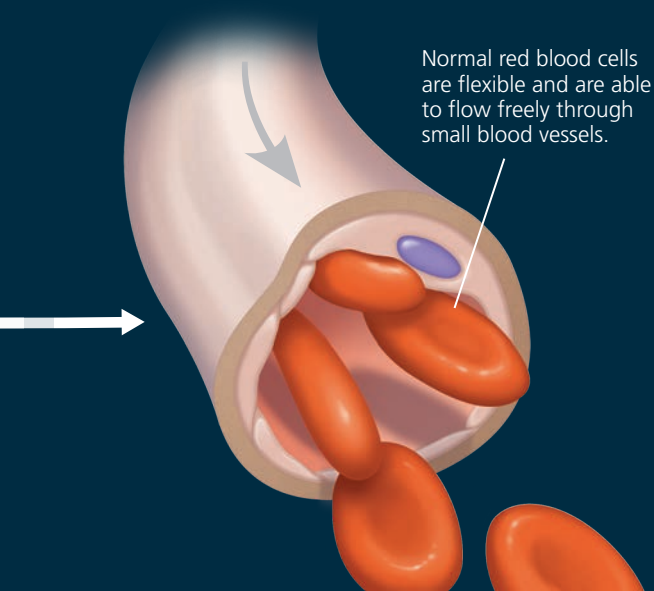
- Homozygotes with two sickle-cell alleles are strongly selected against because of mortality caused by sickle-cell disease. In contrast, heterozygotes experience few harmful effects from sickling yet are more likely to survive malaria than are homozygotes.
- In regions where malaria is common, the net effect of these opposing selective forces is heterozygote advantage. This has caused evolutionary change in populations—the products of which are the areas of relatively high frequencies of the sickle-cell allele shown in the map below.

Effects on Individual Organisms

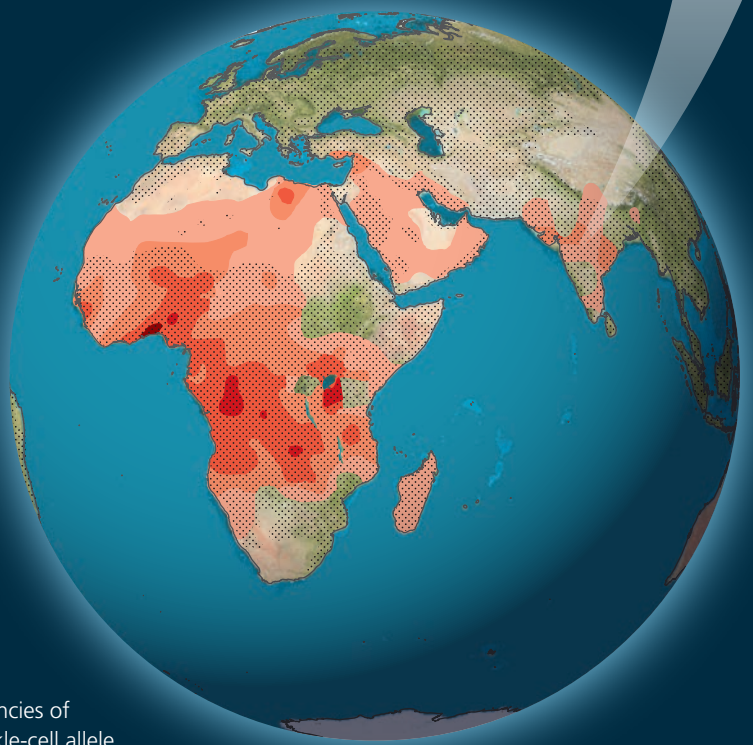
- The formation of sickled red blood cells causes homozygotes with two copies of the sickle-cell allele to have sickle-cell disease.
- Some sickling also occurs in heterozygotes, but not enough to cause the disease; they have sickle-cell trait. (See Figure 14.17.)



The sickled blood cells of a homozygote block small blood vessels, causing great pain and damage to organs such as the heart, kidney, and brain.



Normal red blood cells are flexible and are able to flow freely through small blood vessels.



Key

Frequencies of the sickle-cell allele

- 3.0–6.0%
- 6.0–9.0%
- 9.0–12.0%
- 12.0–15.0%
- >15.0%

Distribution of malaria caused by *Plasmodium falciparum* (a parasitic unicellular eukaryote)

MAKE CONNECTIONS ➤ *In a region free of malaria, would individuals who are heterozygous for the sickle-cell allele be selected for or selected against? Explain.*

4. **Chance, natural selection, and the environment interact.** Chance events can affect the subsequent evolutionary history of populations. For instance, when a storm blows insects or birds hundreds of kilometers over an ocean to an island, the wind does not necessarily transport those individuals that are best suited to the new environment. Thus, not all alleles present in the founding population's gene pool are better suited to the new environment than the alleles that are "left behind." In addition, the environment at a particular location may change unpredictably from year to year, again limiting the extent to which adaptive evolution results in organisms being well suited for current environmental conditions.

With these four constraints, evolution does not tend to craft perfect organisms. Natural selection operates on a "better than" basis. We can, in fact, see evidence for evolution in the many imperfections of the organisms it produces.

CONCEPT CHECK 23.4

1. What is the relative fitness of a sterile mule? Explain.
2. Explain why natural selection is the only evolutionary mechanism that consistently leads to adaptive evolution in a population.
3. **VISUAL SKILLS ➤** Consider a population in which heterozygotes at a certain locus have an extreme phenotype (such as being larger than homozygotes) that confers a selective advantage. Compare this description to the models of selection modes shown in Figure 23.13. Does this situation represent directional, disruptive, or stabilizing selection? Explain your answer.

For suggested answers, see Appendix A.

23 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 23.1

Genetic variation makes evolution possible *(pp. 541–543)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

- **Genetic variation** refers to genetic differences among individuals within a population.
- The nucleotide differences that provide the basis of genetic variation originate when mutation and gene duplication produce new alleles and new genes. New genetic variants are produced rapidly in organisms with short generation times. In sexually reproducing organisms, most of the genetic differences among individuals result from crossing over, the independent assortment of chromosomes, and fertilization.

? *Typically, most of the nucleotide variability that occurs within a genetic locus does not affect the phenotype. Explain why.*

CONCEPT 23.2

The Hardy-Weinberg equation can be used to test whether a population is evolving *(pp. 543–547)*

- A **population**, a localized group of organisms belonging to one species, is united by its **gene pool**, the aggregate of all the alleles in the population.
- For a population in **Hardy-Weinberg equilibrium**, the allele and genotype frequencies will remain constant if the population is large, mating is random, mutation is negligible, there is no gene flow, and there is no natural selection. For such a population, if p and q represent the frequencies of the only two possible alleles at a particular locus, then p^2 is the frequency of one kind of homozygote, q^2 is the frequency of the other kind of homozygote, and $2pq$ is the frequency of the heterozygous genotype.

? *Is it circular reasoning to calculate* p *and* q *from observed genotype frequencies and then use those values of* p *and* q *to test if the population is in Hardy-Weinberg equilibrium? Explain your answer.*

CONCEPT 23.3

Natural selection, genetic drift, and gene flow can alter allele frequencies in a population *(pp. 547–551)*

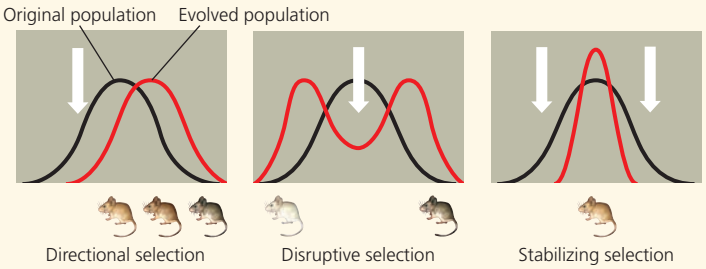
- In natural selection, individuals that have certain inherited traits tend to survive and reproduce at higher rates than other individuals *because of* those traits.
- In **genetic drift**, chance fluctuations in allele frequencies over generations tend to reduce genetic variation.
- **Gene flow**, the transfer of alleles between populations, tends to reduce genetic differences between populations over time.

? *Would two small, geographically isolated populations in very different environments be likely to evolve in similar ways? Explain.*

CONCEPT 23.4

Natural selection is the only mechanism that consistently causes adaptive evolution *(pp. 551–558)*

- One organism has greater **relative fitness** than another organism if it leaves more fertile descendants. The modes of natural selection differ in their effect on phenotype:



- Unlike genetic drift and gene flow, natural selection consistently increases the frequencies of alleles that enhance survival and reproduction, thus improving the degree to which organisms are well-suited for life in their environment.
- **Sexual selection** can result in secondary sex characteristics that can give individuals advantages in mating.

- **Balancing selection** occurs when natural selection maintains two or more forms in a population.
- There are constraints to evolution: Natural selection can act only on available variation; structures result from modified ancestral anatomy; adaptations are often compromises; and chance, natural selection, and the environment interact.

How might secondary sex characteristics in males differ from those in females in a species in which females compete for mates?

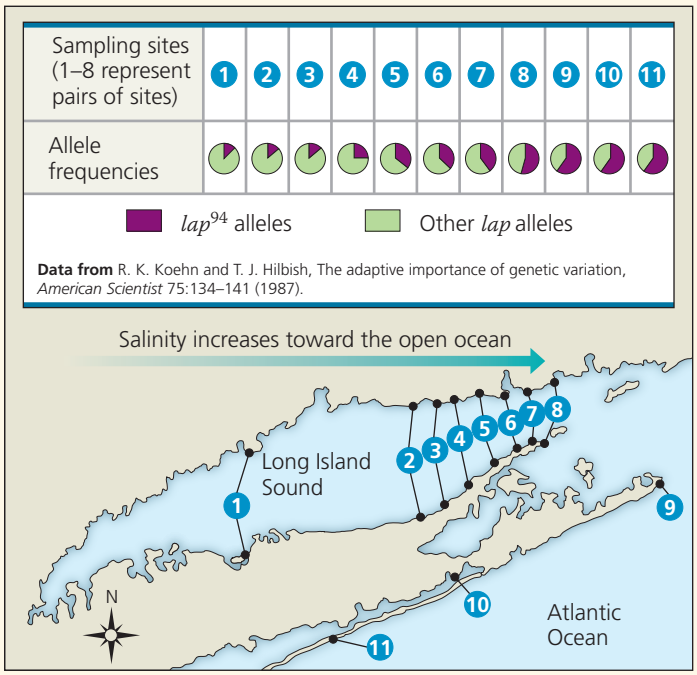
TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–5 can be found in the Study Area in MasteringBiology.

6. **EVOLUTION CONNECTION** Using at least two examples, explain how the process of evolution is revealed by the imperfections of living organisms.

PRACTICE TEST
goo.gl/iAsVgL

7. **SCIENTIFIC INQUIRY • INTERPRET THE DATA** Researchers studied genetic variation in the marine mussel *Mytilus edulis* around Long Island, New York. They measured the frequency of a particular allele (lap^{94}) for an enzyme involved in regulating the mussel's internal saltwater balance. The researchers presented their data as a series of pie charts linked to sampling sites within Long Island Sound, where the salinity is highly variable, and along the coast of the open ocean, where salinity is constant. (a) Create a data table for the 11 sampling sites by estimating the frequency of lap^{94} from the pie charts. (*Hint*: Think of each pie chart as a clock face to help you estimate the proportion of the shaded area.) (b) Graph the frequencies for sites 1–8 to show how the frequency of this allele changes with increasing salinity in Long Island Sound (from southwest to northeast). Evaluate how the data from sites 9–11 compare with the data from the sites within the Sound. (c) Considering the various mechanisms that can alter allele frequency, construct a hypothesis that explains the patterns you observe in the data and that accounts for the following observations: (1) The lap^{94} allele helps mussels maintain osmotic balance in water with a high salt concentration but is costly to use in less salty water; and (2) mussels produce larvae that can disperse long distances before they settle on rocks and grow into adults.



8. **WRITE ABOUT A THEME: ORGANIZATION** Heterozygotes at the sickle-cell locus produce both normal and abnormal (sickle-cell) hemoglobin (see Concept 14.4). When hemoglobin molecules are packed into a heterozygote's red blood cells, some cells receive relatively large quantities of abnormal hemoglobin, making these cells prone to sickling. In a short essay (approximately 100–150 words), explain how these molecular and cellular events lead to emergent properties at the individual and population levels of biological organization.

9. **SYNTHESIZE YOUR KNOWLEDGE**

This kettle lake formed 14,000 years ago when a glacier that covered the surrounding area melted. Initially devoid of animal life, over time the lake was colonized by invertebrates and other animals. Hypothesize how mutation, natural selection, genetic drift, and gene flow may have affected populations that colonized the lake.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

Species and Speciation

24

Figure 24.1 How did this flightless bird come to live on the isolated Galápagos Islands?

KEY CONCEPTS

Galápagos giant tortoise, another species unique to the islands

That "Mystery of Mysteries"

When Darwin came to the Galápagos Islands, he noted that these volcanic islands were teeming with plants and animals found nowhere else in the world **(Figure 24.1)**. Later he realized that these species had formed relatively recently. He wrote in his diary, "Both in space and time, we seem to be brought somewhat near to that great fact—that mystery of mysteries—the first appearance of new beings on this Earth."

The "mystery of mysteries" that captivated Darwin is **speciation**, the process by which one species splits into two or more species. Speciation fascinated Darwin (and many biologists since) because it has produced the tremendous diversity of life, repeatedly yielding new species that differ from existing ones. Later, Darwin realized that speciation also helps to explain the many features that organisms share (the unity of life): When one species splits into two, the species that result share many characteristics because they are descended from this common ancestor. At the DNA sequence level, for example, such similarities indicate that the flightless cormorant (*Phalacrocorax harrisi*) in Figure 24.1 is closely related to flying cormorants found in the Americas. This suggests that the flightless cormorant originated from an ancestral cormorant species that flew from the mainland to the Galápagos.

Speciation also forms a conceptual bridge between **microevolution**, changes over time in allele frequencies in a population, and **macroevolution**, the broad pattern of evolution above the species level. An example of

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

macroevolutionary change is the origin of new groups of organisms, such as mammals or flowering plants, through a series of speciation events. We examined microevolutionary mechanisms in Chapter 23, and we'll turn to macroevolution in Chapter 25. In this chapter, we'll explore the "bridge" between microevolution and macroevolution—the mechanisms by which new species originate from existing ones. First, let's establish what we actually mean by a "species."

HHMI Video: The Origin of Species: The Beak of the Finch

CONCEPT 24.1

The biological species concept emphasizes reproductive isolation

The word *species* is Latin for "kind" or "appearance." In daily life, we commonly distinguish between various "kinds" of organisms—dogs and cats, for instance—based on differences in their appearance. But are organisms truly divided into the discrete units we call species? To answer this question, biologists compare not only the morphology (body form) of different groups of organisms but also less obvious differences in physiology, biochemistry, and DNA sequences. The results generally confirm that morphologically distinct species are indeed discrete groups, differing in many ways besides their body forms.

The Biological Species Concept

The primary definition of species used in this textbook is the **biological species concept**. According to this concept, a **species** is a group of populations whose members have the potential to interbreed in nature and produce viable, fertile offspring—but do not produce viable, fertile offspring with members of other such groups **(Figure 24.2)**. Thus, the members of a biological species are united by being reproductively compatible, at least potentially. All human beings, for example, belong to the same species. A businesswoman in Manhattan may be unlikely to meet a dairy farmer in Mongolia, but if the two should happen to meet and mate, they could have viable babies who develop into fertile adults. In contrast, humans and chimpanzees remain distinct biological species, even where they live in the same region, because many factors keep them from interbreeding and producing fertile offspring.

What holds the gene pool of a species together, causing its members to resemble each other more than they resemble members of other species? Recall the evolutionary mechanism of *gene flow*, the transfer of alleles between populations (see Concept 23.3). Typically, gene flow occurs between the different populations of a species. This ongoing exchange of alleles tends to hold the populations together genetically. But as we'll explore in this chapter, a reduction or lack of gene flow can play a key role in the formation of new species.

(a) Similarity between different species. The eastern meadowlark (*Sturnella magna*, left) and the western meadowlark (*Sturnella neglecta*, right) have similar body shapes and colorations. Nevertheless, they are distinct biological species because their songs and other behaviors are different enough to prevent interbreeding should they meet in the wild.

(b) Diversity within a species. As diverse as we may be in appearance, all humans belong to a single biological species (*Homo sapiens*), defined by our capacity to interbreed successfully.

Figure 24.2 The biological species concept is based on the potential to interbreed, not on physical similarity.

Reproductive Isolation

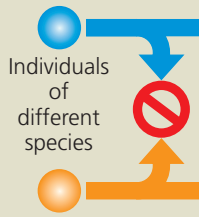
Because biological species are defined in terms of reproductive compatibility, the formation of a new species hinges on **reproductive isolation**—the existence of biological factors (barriers) that impede members of two species from interbreeding and producing viable, fertile offspring. Such barriers block gene flow between the species and limit the formation of **hybrids**, offspring that result from an interspecific mating. Although a single barrier may not prevent all gene flow, a combination of several barriers can effectively isolate a species' gene pool.

Clearly, a fly cannot mate with a frog or a fern, but the reproductive barriers between more closely related species are not so obvious. As described in **Figure 24.3**,

Figure 24.3 Exploring Reproductive Barriers

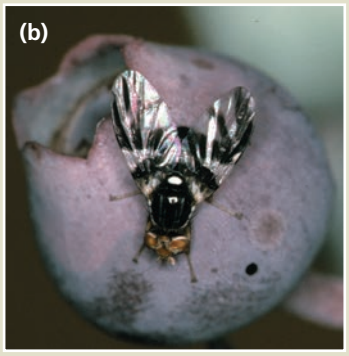
Prezygotic barriers impede mating or hinder fertilization if mating does occur

Habitat Isolation



Two species that occupy different habitats within the same area may encounter each other rarely, if at all, even though they are not isolated by obvious physical barriers, such as mountain ranges.

Example: These two fly species in the genus *Rhagoletis* occur in the same geographic areas, but the apple maggot fly (*Rhagoletis pomonella*) feeds and mates on hawthorns and apples **(a)** while its close relative, the blueberry maggot fly (*R. mendax*), mates and lays its eggs only on blueberries **(b)**.



Temporal Isolation

Species that breed during different times of the day, different seasons, or different years cannot mix their gametes.

Example: In North America, the geographic ranges of the western spotted skunk (*Spilogale gracilis*) **(c)** and the eastern spotted skunk (*Spilogale putorius*) **(d)** overlap, but *S. gracilis* mates in late summer and *S. putorius* mates in late winter.

Behavioral Isolation

Courtship rituals that attract mates and other behaviors unique to a species are effective reproductive barriers, even between closely related species. Such behavioral rituals enable *mate recognition*—a way to identify potential mates of the same species.

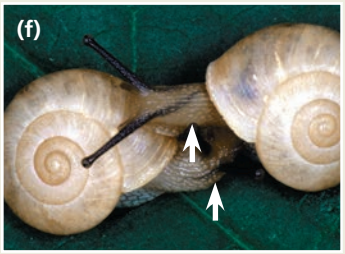
Example: Blue-footed boobies, inhabitants of the Galápagos, mate only after a courtship display unique to their species. Part of the "script" calls for the male to high-step **(e)**, a behavior that calls the female's attention to his bright blue feet.

MB **Video: Blue-Footed Boobies Courtship Ritual**

Mechanical Isolation

Mating is attempted, but morphological differences prevent its successful completion.

Example: The shells of two species of snails in the genus *Bradybaena* spiral in different directions: Moving inward to the center, one spirals in a counterclockwise direction (**f**, left), the other in a clockwise direction (**f**, right). As a result, the snails' genital openings (indicated by arrows) are not aligned, and mating cannot be completed.



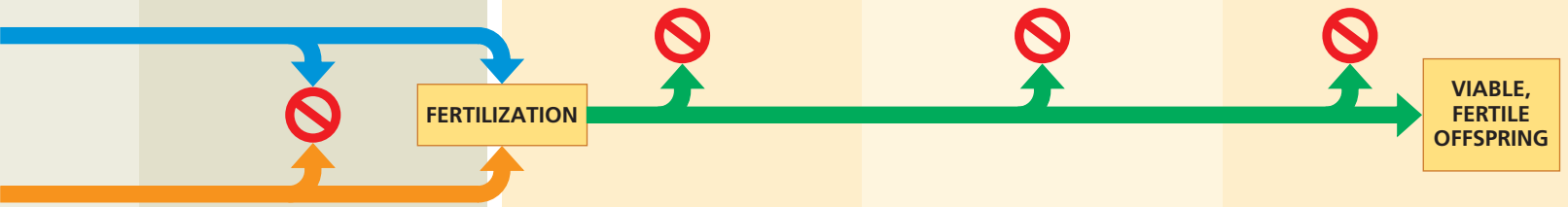
Postzygotic barriers prevent a hybrid zygote from developing into a viable, fertile adult

Gametic Isolation

Reduced Hybrid Viability

Reduced Hybrid Fertility

Hybrid Breakdown



Sperm of one species may not be able to fertilize the eggs of another species. For instance, sperm may not be able to survive in the reproductive tract of females of the other species, or biochemical mechanisms may prevent the sperm from penetrating the membrane surrounding the other species' eggs.

Example: Gametic isolation separates certain closely related species of aquatic animals, such as sea urchins **(g)**. Sea urchins release their sperm and eggs into the surrounding water, where they fuse and form zygotes. It is difficult for gametes of different species, such as the red and purple urchins shown here, to fuse because proteins on the surfaces of the eggs and sperm bind very poorly to each other.

The genes of different parent species may interact in ways that impair the hybrid's development or survival in its environment.

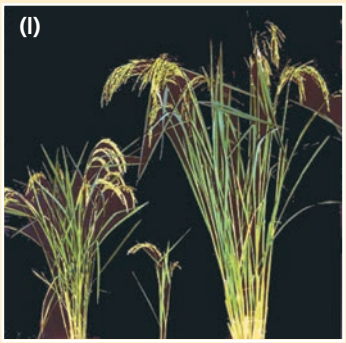
Example: Some salamander subspecies of the genus *Ensatina* live in the same regions and habitats, where they may occasionally hybridize. But most of the hybrids do not complete development, and those that do are frail **(h)**.

Even if hybrids are vigorous, they may be sterile. If the chromosomes of the two parent species differ in number or structure, meiosis in the hybrids may fail to produce normal gametes. Since the infertile hybrids cannot produce offspring when they mate with either parent species, genes cannot flow freely between the species.

Example: The hybrid offspring of a male donkey **(i)** and a female horse **(j)** is a mule **(k)**, which is robust but sterile. A "hinny" (not shown), the offspring of a female donkey and a male horse, is also sterile.

Some first-generation hybrids are viable and fertile, but when they mate with one another or with either parent species, offspring of the next generation are feeble or sterile.

Example: Strains of cultivated rice have accumulated different mutant recessive alleles at two loci in the course of their divergence from a common ancestor. Hybrids between them are vigorous and fertile (**I**, left and right), but plants in the next generation that carry too many of these recessive alleles are small and sterile (**I**, center). Although these rice strains are not yet considered different species, they have begun to be separated by postzygotic barriers.



these barriers can be classified according to whether they contribute to reproductive isolation before or after fertilization. **Prezygotic barriers** ("before the zygote") block fertilization from occurring. Such barriers typically act in one of three ways: by impeding members of different species from attempting to mate, by preventing an attempted mating from being completed successfully, or by hindering fertilization if mating is completed successfully. If a sperm cell from one species overcomes prezygotic barriers and fertilizes an ovum from another species, a variety of **postzygotic barriers** ("after the zygote") may contribute to reproductive isolation after the hybrid zygote is formed. Developmental errors may reduce survival among hybrid embryos. Or problems after birth may cause hybrids to be infertile or decrease their chance of surviving long enough to reproduce.

Limitations of the Biological Species Concept

One strength of the biological species concept is that it directs our attention to a way by which speciation can occur: by the evolution of reproductive isolation. However, the number of species to which this concept can be usefully applied is limited. There is, for example, no way to evaluate the reproductive isolation of fossils. The biological species concept also does not apply to organisms that reproduce asexually all or most of the time, such as prokaryotes. (Many prokaryotes do transfer genes among themselves, as we will discuss in Concept 27.2, but this is not part of their reproductive process.) Furthermore, in the biological species concept, species are designated by the *absence* of gene flow. But there are many pairs of species that are morphologically and ecologically distinct, and yet gene flow occurs between them. An example is the grizzly bear (*Ursus arctos*) and polar bear (*Ursus maritimus*), whose hybrid offspring have been dubbed "grolar bears" **(Figure 24.4)**. As we'll discuss, natural selection can cause such species to remain distinct even though some gene flow occurs between them. Because of the limitations to the biological species concept, alternative species concepts are useful in certain situations.

Other Definitions of Species

While the biological species concept emphasizes the *separateness* of different species due to reproductive barriers, several other definitions emphasize the *unity within* a species. For example, the **morphological species concept** distinguishes a species by body shape and other structural features. The morphological species concept can be applied to asexual and sexual organisms, and it can be useful even without information on the extent of gene flow. In practice, scientists often distinguish species using morphological criteria. A disadvantage of this approach, however, is that it relies on subjective criteria; researchers may disagree on which structural features distinguish a species.

▲ Figure 24.4 Hybridization between two species of bears in the genus *Ursus*.

The **ecological species concept** defines a species in terms of its ecological niche, the sum of how members of the species interact with the nonliving and living parts of their environment (see Concept 54.1). For example, two species of oak trees might differ in their size or in their ability to tolerate dry conditions, yet still occasionally interbreed. Because they occupy different ecological niches, these oaks would be considered separate species even though they are connected by some gene flow. Unlike the biological species concept, the ecological species concept can accommodate asexual as well as sexual species. It also emphasizes the role of disruptive natural selection as organisms adapt to different environments.

In addition to those discussed here, more than 20 other species definitions have been proposed. The usefulness of each definition depends on the situation and the research questions being asked. For our purposes of studying how species originate, the biological species concept, with its focus on reproductive barriers, is particularly helpful.

CONCEPT CHECK 24.1

1. (a) Which species concept(s) could you apply to both asexual and sexual species? (b) Which would be most useful for identifying species in the field? Explain.
2. **WHAT IF? ➤** Suppose two bird species live in a forest and are not known to interbreed. One species feeds and mates in the treetops and the other on the ground. But in captivity, the birds can interbreed and produce viable, fertile offspring. What type of reproductive barrier most likely keeps these species separate in nature? Explain.

For suggested answers, see Appendix A.

CONCEPT 24.2

Speciation can take place with or without geographic separation

Having discussed what constitutes a unique species, let's return to the process by which such species arise from existing species. We'll describe this process by focusing on the geographic setting in which gene flow is interrupted between populations of the existing species—in allopatric speciation the populations are geographically isolated, while in sympatric speciation they are not **(Figure 24.5)**.

Allopatric ("Other Country") Speciation

In **allopatric speciation** (from the Greek *allos*, other, and *patra*, homeland), gene flow is interrupted when a population is divided into geographically isolated subpopulations. For example, the water level in a lake may subside, resulting in two or more smaller lakes that are now home to separated populations (see Figure 24.5a). Or a river may change course and divide a population of animals that cannot cross it. Allopatric speciation can also occur without geologic remodeling, such as when individuals colonize a remote area and their descendants become geographically isolated from the parent population. The flightless cormorant shown in Figure 24.1 probably originated in this way from an ancestral flying species that reached the Galápagos Islands.

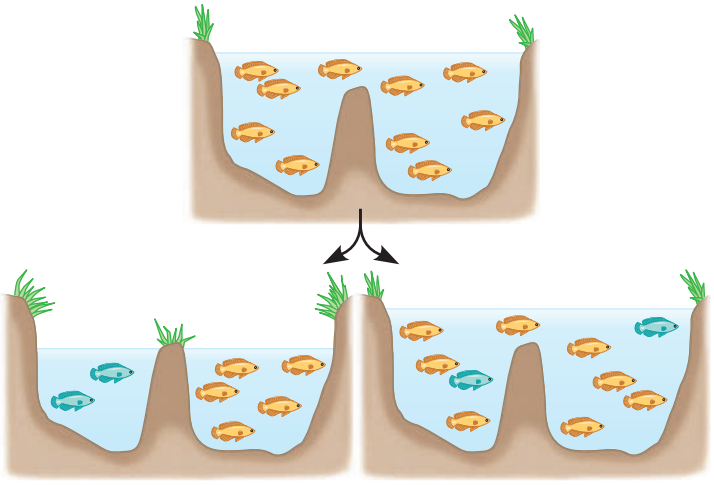
The Process of Allopatric Speciation

How formidable must a geographic barrier be to promote allopatric speciation? The answer depends on the ability of the organisms to move about. Birds, mountain lions, and coyotes can cross rivers and canyons—as can the windblown pollen of pine trees and the seeds of some flowering plants. In contrast, small rodents may find a wide river or deep canyon a formidable barrier.

Once geographic isolation has occurred, the separated gene pools may diverge. Different mutations arise, and natural selection and genetic drift may alter allele frequencies in different ways in the separated populations. Reproductive isolation may then evolve as a by-product of the genetic divergence that results from selection or drift.

Figure 24.6 describes an example. On Andros Island, in the Bahamas, populations of the mosquitofish *Gambusia hubbsi* colonized a series of ponds that later became isolated from one another. Genetic analyses indicate that little or no gene flow currently occurs between the ponds. The environments of these ponds are very similar except that some contain predatory fishes, while others do not. In ponds with predatory fishes, selection has favored the evolution of a mosquitofish body shape that enables rapid bursts of speed (Figure 24.6). In ponds without predatory fishes, selection has favored a different body shape, one that improves the ability to swim for long periods of time. How have these different selective pressures affected the evolution of reproductive barriers? Researchers studied this question by bringing together mosquitofish from the two types of ponds. They found that female mosquitofish prefer to mate with males whose body shape is similar to their own. This preference establishes a behavioral barrier to reproduction between mosquitofish from ponds with predators and those

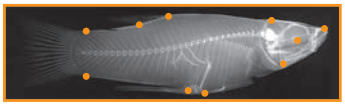
Figure 24.5 The geography of speciation.



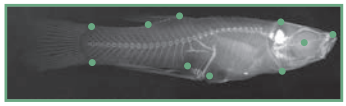
(a) Allopatric speciation. A population forms a new species while geographically isolated from its parent population.

(b) Sympatric speciation. A subset of a population forms a new species without geographic separation.

Figure 24.6 Evolution in mosquitofish populations. Different body shapes have evolved in mosquitofish populations from ponds with and without predators. These differences affect how quickly the fish can accelerate to escape and their survival rate when exposed to predators.

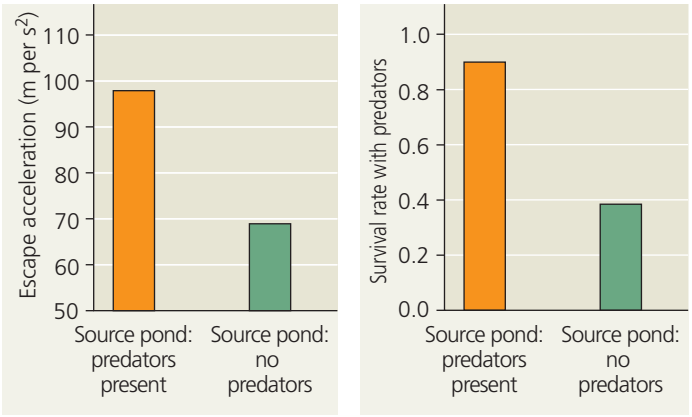


In ponds with predatory fishes, the mosquitofish's head is streamlined and the tail is powerful, enabling rapid bursts of speed.



In ponds without predatory fishes, mosquitofish have a different body shape that favors long, steady swimming.

(a) Differences in body shape



(b) Differences in escape acceleration and survival

from ponds without predators. Thus, as a by-product of selection for avoiding predators, reproductive barriers have formed in these allopatric populations.

Evidence of Allopatric Speciation

Many studies provide evidence that speciation can occur in allopatric populations. For example, laboratory studies show that reproductive barriers can develop when populations are isolated experimentally and subjected to different environmental conditions **(Figure 24.7)**.

Field studies indicate that allopatric speciation also can occur in nature. Consider the 30 species of snapping shrimp in the genus *Alpheus* that live off the Isthmus of Panama, the land bridge that connects South and North America **(Figure 24.8)**. Fifteen of these species live on the Atlantic side of the isthmus, while the other 15 live on the Pacific side. Before the isthmus formed, gene flow could occur between the Atlantic and Pacific populations of snapping shrimp. Did the species on different sides of the isthmus originate by allopatric speciation? Morphological and genetic data group these shrimp into 15 pairs of *sister species*, pairs whose member species are each other's closest relative.

Figure 24.8 Allopatric speciation in snapping shrimp (*Alpheus*). The shrimp pictured are just 2 of the 15 pairs of sister species that arose as populations were divided by the formation of the Isthmus of Panama. The color-coded type indicates the sister species.

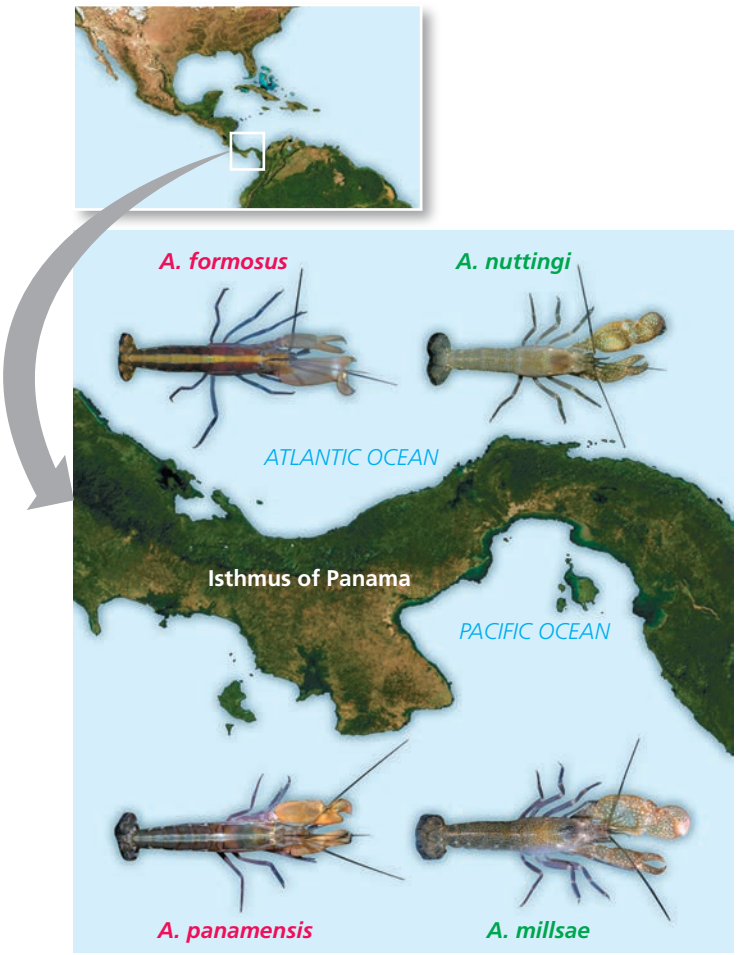
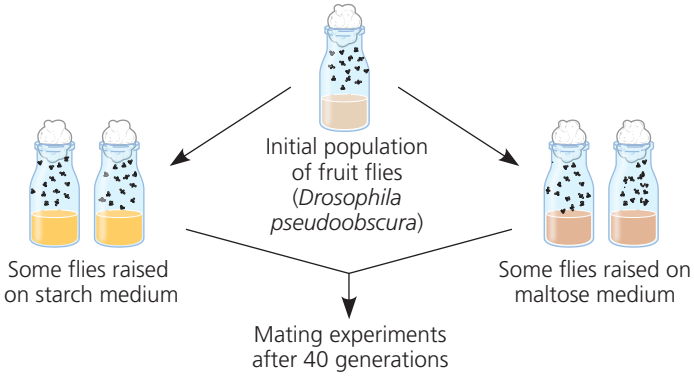


Figure 24.7

Inquiry Can divergence of allopatric populations lead to reproductive isolation?

Experiment A researcher divided a laboratory population of the fruit fly *Drosophila pseudoobscura*, raising some flies on a starch medium and others on a maltose medium. After one year (about 40 generations), natural selection resulted in divergent evolution: Populations raised on starch digested starch more efficiently, while those raised on maltose digested maltose more efficiently. The researcher then put flies from the same or different populations in mating cages and measured mating frequencies. All flies used in the mating preference tests were reared for one generation on a standard cornmeal medium.



Results Mating patterns among populations of flies raised on different media are shown below. When flies from "starch populations" were mixed with flies from "maltose populations," the flies tended to mate with like partners. But in the control group (shown on the right), flies from different populations adapted to starch were about as likely to mate with each other as with flies from their own population; similar results were obtained for control groups adapted to maltose.

		Female	
		Starch	Maltose
Male	Starch	22	9
	Maltose	8	20

Number of matings in experimental group

		Female	
		Starch population 1	Starch population 2
Male	Starch population 1	18	15
	Starch population 2	12	15

Number of matings in control group

Conclusion In the experimental group, the strong preference of "starch flies" and "maltose flies" to mate with like-adapted flies indicates that a reproductive barrier was forming between these fly populations. Although this barrier was not absolute (some mating between starch flies and maltose flies did occur), after 40 generations reproductive isolation appeared to be increasing. This barrier may have been caused by differences in courtship behavior that arose as an incidental by-product of differing selective pressures as these allopatric populations adapted to different sources of food.

Data from D. M. B. Dodd, Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*, *Evolution* 43:1308–1311 (1989).

WHAT IF? ➤ *Why were all flies used in the mating preference tests reared on a standard medium (rather than on starch or maltose)?*

In each of these 15 pairs, one of the sister species lives on the Atlantic side of the isthmus, while the other lives on the Pacific side. This fact strongly suggests that the two species arose as a consequence of geographic separation. Furthermore, genetic analyses indicate that the *Alpheus* species originated from 9 to 3 million years ago, with the sister species that live in the deepest water diverging first. These divergence times are consistent with geologic evidence that the isthmus formed gradually, starting 10 million years ago, and closing completely about 3 million years ago.

The importance of allopatric speciation is also suggested by the fact that regions that are isolated or highly subdivided by barriers typically have more species than do otherwise similar regions that lack such features. For example, many unique plants and animals are found on the geographically isolated Hawaiian Islands (we'll return to the origin of Hawaiian species in Concept 25.4). Field studies also show that reproductive isolation between two populations generally increases as the geographic distance between them increases, a finding consistent with allopatric speciation. In the **Scientific Skills Exercise**, you will analyze data from one such study that examined reproductive isolation in geographically separated salamander populations.

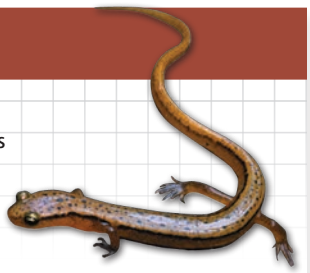
Note that while geographic isolation prevents interbreeding between members of allopatric populations, physical separation is not a biological barrier to reproduction. Biological reproductive barriers such as those described in Figure 24.3 are intrinsic to the organisms themselves. Hence, it is biological barriers that can prevent interbreeding when members of different populations come into contact with one another.

Sympatric ("Same Country") Speciation

In **sympatric speciation** (from the Greek *syn*, together), speciation occurs in populations that live in the same geographic area (see Figure 24.5b). How can reproductive barriers form between sympatric populations while their members remain in contact with each other? Although such contact (and the ongoing gene flow that results) makes sympatric speciation less common than allopatric speciation, sympatric speciation can occur if gene flow is reduced by such factors as polyploidy, sexual selection, and habitat differentiation. (Note that these factors can also promote allopatric speciation.)

SCIENTIFIC SKILLS EXERCISE

Identifying Independent and Dependent Variables, Making a Scatter Plot, and Interpreting Data



Does Distance Between Salamander Populations Increase Their Reproductive Isolation? Allopatric speciation begins when populations become geographically isolated, preventing mating between individuals in different populations and thus stopping gene flow. It is logical that as distance between populations increases, so will their degree of reproductive isolation. To test this hypothesis, researchers studied populations of the dusky salamander (*Desmognathus ochrophaeus*) living on different mountain ranges in the southern Appalachians.

How the Experiment Was Done The researchers tested the reproductive isolation of pairs of salamander populations by leaving one male and one female together and later checking the females for the presence of sperm. Four mating combinations were tested for each pair of populations (A and B)—two *within* the same population (female A with male A and female B with male B) and two *between* populations (female A with male B and female B with male A).

Data from the Experiment The researchers used an index of reproductive isolation that ranged from a value of 0 (no isolation) to a value of 2 (full isolation). The proportion of successful matings for each mating combination was measured, with 100% success = 1 and no success = 0. The reproductive isolation value for two populations is the sum of the proportion of successful matings of each type within populations (AA + BB) minus the sum of the proportion of successful matings of each type between populations (AB + BA). The table provides distance and reproductive isolation data for 27 pairs of dusky salamander populations.

INTERPRET THE DATA

1. State the researchers' hypothesis, and identify the independent and dependent variables in this study. Explain why the researchers used four mating combinations for each pair of populations.
2. Calculate the value of the reproductive isolation index if (a) *all* of the matings within a population were successful, but *none* of the matings between populations were successful; (b) salamanders are equally successful in mating with members of their own population and members of another population.
3. Make a scatter plot to help you visualize any patterns that might indicate a relationship between the variables. Plot the independent variable on the *x*-axis and the dependent variable on the *y*-axis. (For additional information about graphs, see the Scientific Skills Review in Appendix F and the Study Area of MasteringBiology.)
4. Interpret your graph by (a) explaining in words any pattern indicating a possible relationship between the variables and (b) hypothesizing the possible cause of such a relationship.

MB **Instructors**: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

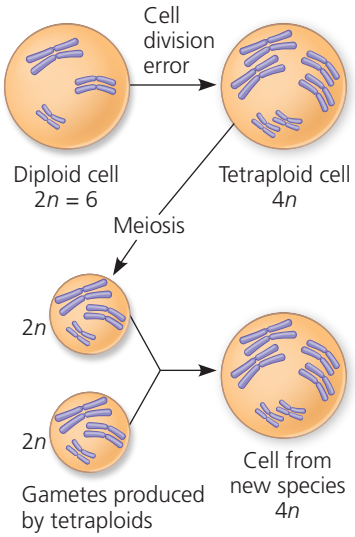
Data from S. G. Tilley, A. Verrell, and S. J. Arnold, Correspondence between sexual isolation and allozyme differentiation: a test in the salamander *Desmognathus ochrophaeus*, *Proceedings of the National Academy of Sciences USA* 87:2715–2719 (1990).

Geographic Distance (km)	15	32	40	47	42	62	63	81	86	107	107	115	137	147
Reproductive Isolation Value	0.32	0.54	0.50	0.50	0.82	0.37	0.67	0.53	1.15	0.73	0.82	0.81	0.87	0.87
Distance (continued)	137	150	165	189	219	239	247	53	55	62	105	179	169	
Isolation (continued)	0.50	0.57	0.91	0.93	1.5	1.22	0.82	0.99	0.21	0.56	0.41	0.72	1.15	

Polyploidy

A species may originate from an accident during cell division that results in extra sets of chromosomes, a condition called **polyploidy**. Polyploid speciation occasionally occurs in animals; for example, the gray tree frog *Hyla versicolor* (see Figure 23.16) is thought to have originated in this way. However, polyploidy is far more common in plants. In fact, botanists estimate that more than 80% of the plant species alive today are descended from ancestors that formed by polyploid speciation.

Figure 24.9 Sympatric speciation by autopolyploidy.



Two distinct forms of polyploidy have been observed in plant (and a few animal) populations. An **autopolyploid** (from the Greek *autos*, self) is an individual that has more than two chromosome sets that are all derived from a single species. In plants, for example, a failure of cell division could double a cell's chromosome number from the original number ($2n$) to a tetraploid number ($4n$) **(Figure 24.9)**.

A tetraploid can produce fertile tetraploid offspring by self-pollinating or by mating with other tetraploids. In addition, the tetraploids are reproductively isolated from $2n$ plants of the original population, because the triploid ($3n$) offspring of such unions have reduced fertility. Thus, in just one generation, autopolyploidy can generate reproductive isolation without any geographic separation.

A second form of polyploidy can occur when two different species interbreed and produce hybrid offspring. Most such hybrids are sterile because the set of chromosomes from one species cannot pair during meiosis with the set of chromosomes from the other species. However, an infertile hybrid may be able to propagate itself asexually (as many plants can do). In subsequent generations, various mechanisms can change a sterile hybrid into a fertile polyploid called an **allopolyploid** **(Figure 24.10)**. The allopolyploids are fertile when mating with each other but cannot interbreed with either parent species; thus, they represent a new biological species.

Although it can be challenging to study speciation in the field, scientists have documented at least five new plant species that have originated by polyploid speciation since 1850. One of these examples involves the origin of a new species of goatsbeard plant (genus *Tragopogon*) in the Pacific Northwest. *Tragopogon* first arrived in the region when humans introduced three European species in the early 1900s: *T. pratensis, T. dubius,* and *T. porrifolius.* These three species are now common weeds

Figure 24.10 One mechanism for allopolyploid speciation in plants. Most hybrids are sterile because their chromosomes are not homologous and cannot pair during meiosis. However, such a hybrid may be able to reproduce asexually. This diagram traces one mechanism that can produce fertile hybrids (allopolyploids) that are members of a new species. The new species has a diploid chromosome number equal to the sum of the diploid chromosome numbers of the two parent species.

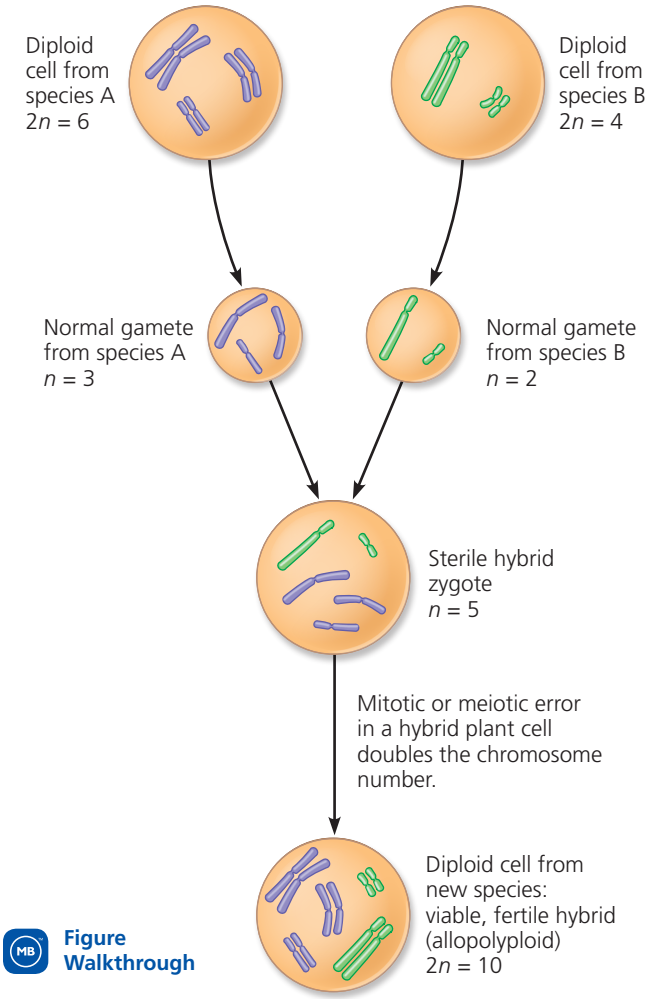
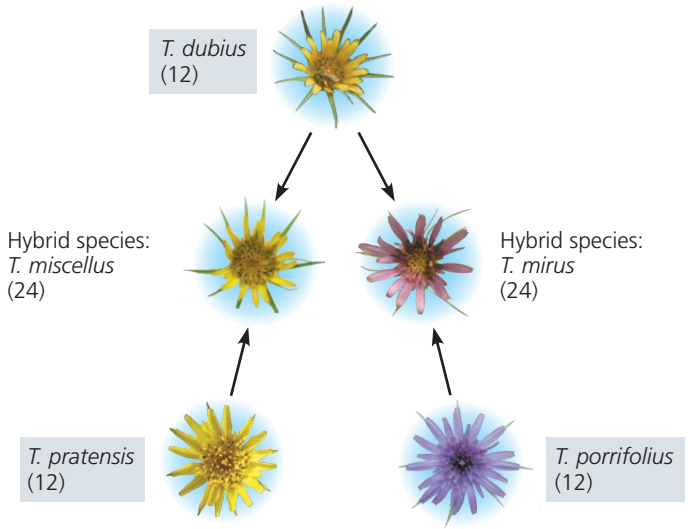


Figure Walkthrough

in abandoned parking lots and other urban sites. In 1950, a new *Tragopogon* species was discovered near the Idaho-Washington border, a region where all three European species also were found. Genetic analyses revealed that this new species, *Tragopogon miscellus*, is a hybrid of two of the European species **(Figure 24.11)**. Although the *T. miscellus* population grows mainly by reproduction of its own members, additional episodes of hybridization between the parent species continue to add new members to the *T. miscellus* population. Later, scientists discovered another new *Tragopogon* species, *T. mirus*—this one a hybrid of *T. dubius* and *T. porrifolius* (see Figure 24.11). The *Tragopogon* story is just one of several well-studied examples in which scientists have observed speciation in progress.

Many important agricultural crops—such as oats, cotton, potatoes, tobacco, and wheat—are polyploids. For example, the wheat used for bread, *Triticum aestivum*, is an allohexaploid (six sets of chromosomes, two sets from each of three different species). The first of the polyploidy events that eventually led

Figure 24.11 Allopolyploid speciation in *Tragopogon*. The gray boxes indicate the three parent species. The diploid chromosome number of each species is shown in parentheses.



to modern wheat probably occurred about 8,000 years ago in the Middle East as a spontaneous hybrid of an early cultivated wheat species and a wild grass. Today, plant geneticists generate new polyploids in the laboratory by using chemicals that induce meiotic and mitotic errors. By harnessing the evolutionary process, researchers can produce new hybrid species with desired qualities, such as a hybrid that combines the high yield of wheat with the hardiness of rye.

Animation: Speciation by Changes in Ploidy

Sexual Selection

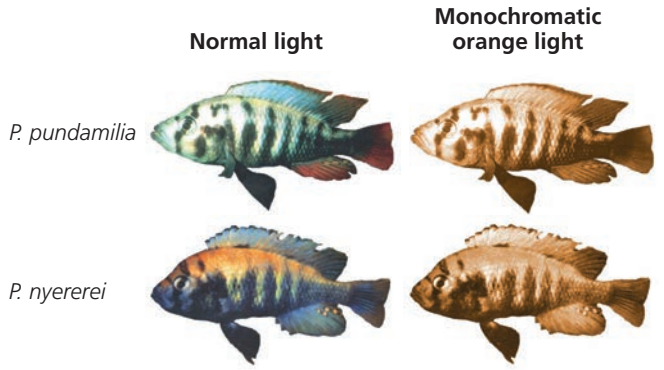
There is evidence that sympatric speciation can also be driven by sexual selection. Clues to how this can occur have been found in cichlid fish from one of Earth's hot spots of animal speciation, East Africa's Lake Victoria. This lake was once home to as many as 600 species of cichlids. Genetic data indicate that these species originated within the last 100,000 years from a small number of colonizing species that arrived from other lakes and rivers. How did so many species—more than double the number of freshwater fish species known in all of Europe—originate within a single lake?

One hypothesis is that subgroups of the original cichlid populations adapted to different food sources and the resulting genetic divergence contributed to speciation in Lake Victoria. But sexual selection, in which (typically) females select males based on their appearance (see Concept 23.4), may also have been a factor. Researchers have studied two closely related sympatric species of cichlids that differ mainly in the coloration of breeding males: Breeding *Pundamilia pundamilia* males have a blue-tinged back, whereas breeding *Pundamilia nyererei* males have a red-tinged back **(Figure 24.12)**. Their results suggest that mate choice based on male breeding coloration can act as a reproductive barrier that keeps the gene pools of these two species separate.

Figure 24.12

Inquiry Does sexual selection in cichlids result in reproductive isolation?

Experiment Researchers placed males and females of *Pundamilia pundamilia* and *P. nyererei* together in two aquarium tanks, one with natural light and one with a monochromatic orange lamp. Under normal light, the two species are noticeably different in male breeding coloration; under monochromatic orange light, the two species are very similar in color. The researchers then observed the mate choices of the females in each tank.



Results Under normal light, females of each species strongly preferred males of their own species. But under orange light, females of each species responded indiscriminately to males of both species. The resulting hybrids were viable and fertile.

Conclusion The researchers concluded that mate choice by females based on male breeding coloration can act as a reproductive barrier that keeps the gene pools of these two species separate. Since the species can still interbreed when this prezygotic behavioral barrier is breached in the laboratory, the genetic divergence between the species is likely to be small. This suggests that speciation in nature has occurred relatively recently.

Data from O. Seehausen and J. J. M. van Alphen, The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex), *Behavioral Ecology and Sociobiology* 42:1–8 (1998).

WHAT IF? ➤ *Suppose that female cichlids living in the murky waters of a polluted lake could not distinguish colors well. In such waters, how might the gene pools of these species change over time?*

Habitat Differentiation

Sympatric speciation can also occur when a subpopulation exploits a habitat or resource not used by the parent population. Consider the North American apple maggot fly (*Rhagoletis pomonella*), a pest of apples. The fly's original habitat was the native hawthorn tree (see Figure 24.3a), but about 200 years ago, some populations colonized apple trees that had been introduced by European settlers. Apple maggot flies usually mate on or near their host plant. This results in a prezygotic barrier (habitat isolation) between populations that feed on apples and populations that feed on hawthorns. Furthermore, because apples mature more quickly than hawthorn fruit, natural selection has favored apple-feeding flies with rapid development. These apple-feeding populations now show temporal isolation from the hawthorn-feeding *R. pomonella*, providing a second prezygotic barrier to gene flow between

the two populations. Researchers also have identified alleles that benefit the flies that use one host plant but harm the flies that use the other host plant. Natural selection operating on these alleles has provided a postzygotic barrier to reproduction, further limiting gene flow. Altogether, although the two populations are still classified as subspecies rather than separate species, sympatric speciation appears to be well under way.

Allopatric and Sympatric Speciation: *A Review*

Now let's recap the processes by which new species form. In allopatric speciation, a new species forms in geographic isolation from its parent population. Geographic isolation severely restricts gene flow. Intrinsic barriers to reproduction with members of the parent population may then arise as a by-product of genetic changes that occur within the isolated population. Many different processes can produce such genetic changes, including natural selection under different environmental conditions, genetic drift, and sexual selection. Once formed, reproductive barriers that arise in allopatric populations can prevent interbreeding with the parent population even if the populations come back into contact.

Sympatric speciation, in contrast, requires the emergence of a reproductive barrier that isolates a subset of a population from the remainder of the population in the same area. Though rarer than allopatric speciation, sympatric speciation can occur when gene flow to and from the isolated subpopulation is blocked. This can occur as a result of polyploidy, a condition in which an organism has extra sets of chromosomes. Sympatric speciation also can result from sexual selection. Finally, sympatric speciation can occur when a subset of a population becomes reproductively isolated because of natural selection that results from a switch to a habitat or food source not used by the parent population.

Having reviewed the geographic context in which species originate, we'll next explore in more detail what can happen when new or partially formed species come into contact.

HHMI Video: Anole Lizards: An Example of Speciation

CONCEPT CHECK 24.2

1. Summarize key differences between allopatric and sympatric speciation. Which type of speciation is more common, and why?
2. Describe two mechanisms that can decrease gene flow in sympatric populations, thereby making sympatric speciation more likely to occur.
3. **WHAT IF? ➤** Is allopatric speciation more likely to occur on an island close to a mainland or on a more isolated island of the same size? Explain your prediction.
4. **MAKE CONNECTIONS ➤** Review the process of meiosis in Figure 13.8. Describe how an error during meiosis could lead to polyploidy.

For suggested answers, see Appendix A.

CONCEPT 24.3

Hybrid zones reveal factors that cause reproductive isolation

What happens if species with incomplete reproductive barriers come into contact with one another? One possible outcome is the formation of a **hybrid zone**, a region in which members of different species meet and mate, producing at least some offspring of mixed ancestry. In this section, we'll explore hybrid zones and what they reveal about factors that cause the evolution of reproductive isolation.

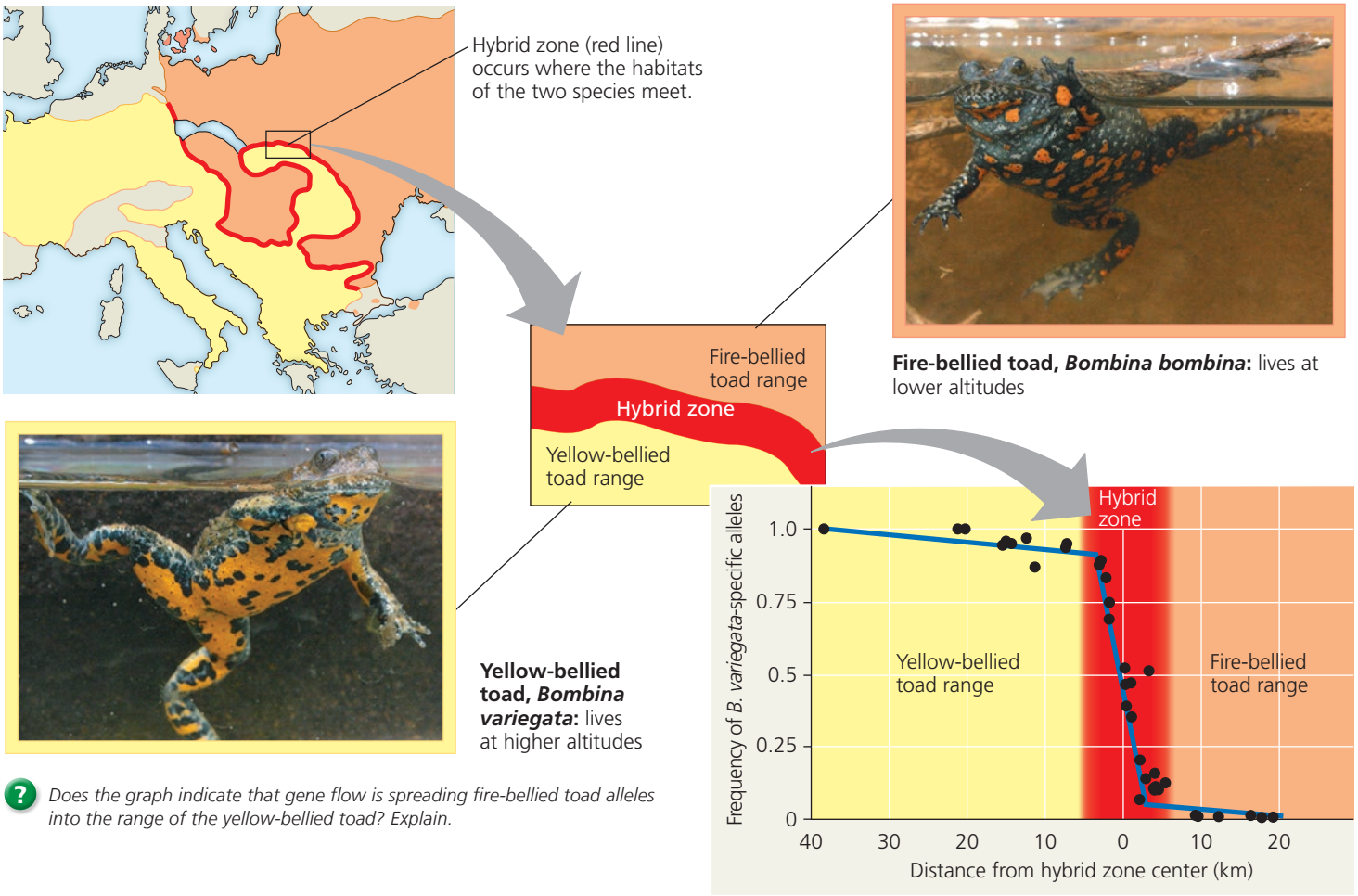
Patterns Within Hybrid Zones

Some hybrid zones form as narrow bands, such as the one depicted in **Figure 24.13** for the yellow-bellied toad (*Bombina variegata*) and its close relative, the fire-bellied toad (*B. bombina*). This hybrid zone, represented by the red line on the map, extends for 4,000 km but is less than 10 km wide in most places. The hybrid zone occurs where the higher-altitude habitat of the yellow-bellied toad meets the lowland habitat of the fire-bellied toad. Across a given "slice" of the zone, the frequency of alleles specific to yellow-bellied toads typically decreases from close to 100% at the edge where only yellow-bellied toads are found to around 50% in the central portion of the zone to close to 0% at the edge where only fire-bellied toads are found.

What causes such a pattern of allele frequencies across a hybrid zone? We can infer that there is an obstacle to gene flow—otherwise, alleles from one parent species would also be common in the gene pool of the other parent species. Are geographic barriers reducing gene flow? Not in this case, since the toads can move throughout the hybrid zone. A more important factor is that hybrid toads have increased rates of embryonic mortality and a variety of morphological abnormalities, including ribs that are fused to the spine and malformed tadpole mouthparts. Because the hybrids have poor survival and reproduction, they produce few viable offspring with members of the parent species. As a result, hybrid individuals rarely serve as a stepping-stone from which alleles are passed from one species to the other. Outside the hybrid zone, additional obstacles to gene flow may be provided by natural selection in the different environments in which the parent species live.

Hybrid zones typically are located wherever the habitats of the interbreeding species meet. Those regions often resemble a group of isolated patches scattered across the landscape—more like the complex pattern of spots on a Dalmatian than the continuous band shown in Figure 24.13. But regardless of whether they have complex or simple spatial patterns, hybrid zones form when two species lacking complete barriers to reproduction come into contact. What happens when the habitats of the interbreeding species change over time?

Figure 24.13 A narrow hybrid zone for *Bombina* toads in Europe. The graph shows species-specific allele frequencies across the width of the zone near Krakow, Poland, averaged over six genetic loci. A value of 1.0 indicates that all individuals were yellow-bellied toads, 0 indicates that all individuals were fire-bellied toads, and intermediate frequencies indicate that some individuals were of mixed ancestry.



? *Does the graph indicate that gene flow is spreading fire-bellied toad alleles into the range of the yellow-bellied toad? Explain.*

Hybrid Zones and Environmental Change

A change in environmental conditions can alter where the habitats of interbreeding species meet. When this happens, an existing hybrid zone can move to a new location, or a novel hybrid zone may form.

For example, black-capped chickadees (*Poecile atricapillus*) and Carolina chickadees (*P. carolinensis*) interbreed in a narrow hybrid zone that runs from New Jersey to Kansas. Recent studies have shown that the location of this hybrid zone has shifted northward as the climate has warmed. In another example, a series of warm winters prior to 2003 enabled the southern flying squirrel (*Glaucomys volans*) to expand northward into the range of the northern flying squirrel, *G. sabrinus*. Previously, the ranges of these two species had not overlapped. Genetic analyses showed that these flying squirrels began to hybridize where their ranges came into contact, thereby forming a novel hybrid zone induced by climate change.

Finally, note that a hybrid zone can be a source of novel genetic variation that improves the ability of one or both parent species to cope with changing environmental conditions. This can occur when an allele found only in one parent species is transferred first to hybrid individuals, and then to the other parent species when hybrids breed with the second parent species. Recent genetic analyses have shown that hybridization has been a source for such novel genetic variation in various insect, bird, and plant species. In the **Problem-Solving Exercise**, you can examine one such example: a case in which hybridization may have led to the transfer of insecticide-resistance alleles between mosquitoes that transmit malaria.

Hybrid Zones over Time

Studying a hybrid zone is like observing a naturally occurring experiment on speciation. Will the hybrids become reproductively isolated from their parents and form a new species, as occurred by polyploidy in the goatsbeard plant of the Pacific Northwest? If not, there are three other common outcomes for the hybrid zone over time: reinforcement of

PROBLEM-SOLVING EXERCISE

Is hybridization promoting insecticide resistance in mosquitoes that transmit malaria?

Malaria is a leading cause of human illness and mortality worldwide, with 200 million people infected and 600,000 deaths each year. In the 1960s, the incidence of malaria was reduced owing to the use of insecticides that killed mosquitoes in the genus *Anopheles*, which transmit the disease from person to person. But today, mosquitoes are becoming resistant to insecticides—causing a resurgence in malaria.

Insecticide-treated bed nets have helped reduce cases of malaria in many countries, but resistance to insecticides is rising in mosquito populations.

Instructors: A version of this Problem-Solving Exercise can be assigned in MasteringBiology.

In this exercise, you will investigate whether alleles encoding resistance to insecticides have been transferred between closely related species of *Anopheles*.

Your Approach The principle guiding your investigation is that DNA analyses can detect the transfer of resistance alleles between closely related mosquito species. To find out whether such transfers have occurred, you will analyze DNA results from two species of mosquitoes that transmit malaria (*Anopheles gambiae* and *A. coluzzii*) and from *A. gambiae* × *A. coluzzii* hybrids.

Your Data Resistance to DDT and other insecticides in *Anopheles* is affected by a sodium channel gene, *kdr*. The *r* allele of this gene confers resistance, while the wild type (+/+) genotype is not resistant. Researchers sequenced the *kdr* gene from mosquitoes collected in Mali during three time periods: pre-2006 (2002 and 2004), 2006, and post-2006 (2009–2012). *A. gambiae* and *A. coluzzii* were collected during all three time periods, but their hybrids only occurred in 2006, the first year that insecticide-treated bed nets were used to reduce the spread of malaria. A likely explanation is that the introduction of the treated bed nets may have briefly favored hybrid individuals, which are usually at a selective disadvantage.

Observed numbers of mosquitoes by *kdr* genotype			
	+/+	+/*r*	*r*/*r*
A. gambiae			
Pre-2006	3	5	2
2006	8	8	7
Post-2006	3	3	57
Hybrids			
2006	10	7	0
A. coluzzii			
Pre-2006	226	0	0
2006	70	7	0
Post-2006	79	127	94

Your Analysis

1. How did the frequencies of *kdr* genotypes change over time in *A. gambiae*? Describe a hypothesis that accounts for these observations.
2. How did the frequencies of *kdr* genotypes change over time in *A. coluzzii*? Describe a hypothesis that accounts for these observations.
3. Do these results indicate that hybridization can lead to the transfer of adaptive alleles? Explain.
4. Predict how the transfer of the *r* allele to *A. coluzzii* populations could affect the number of malaria cases.

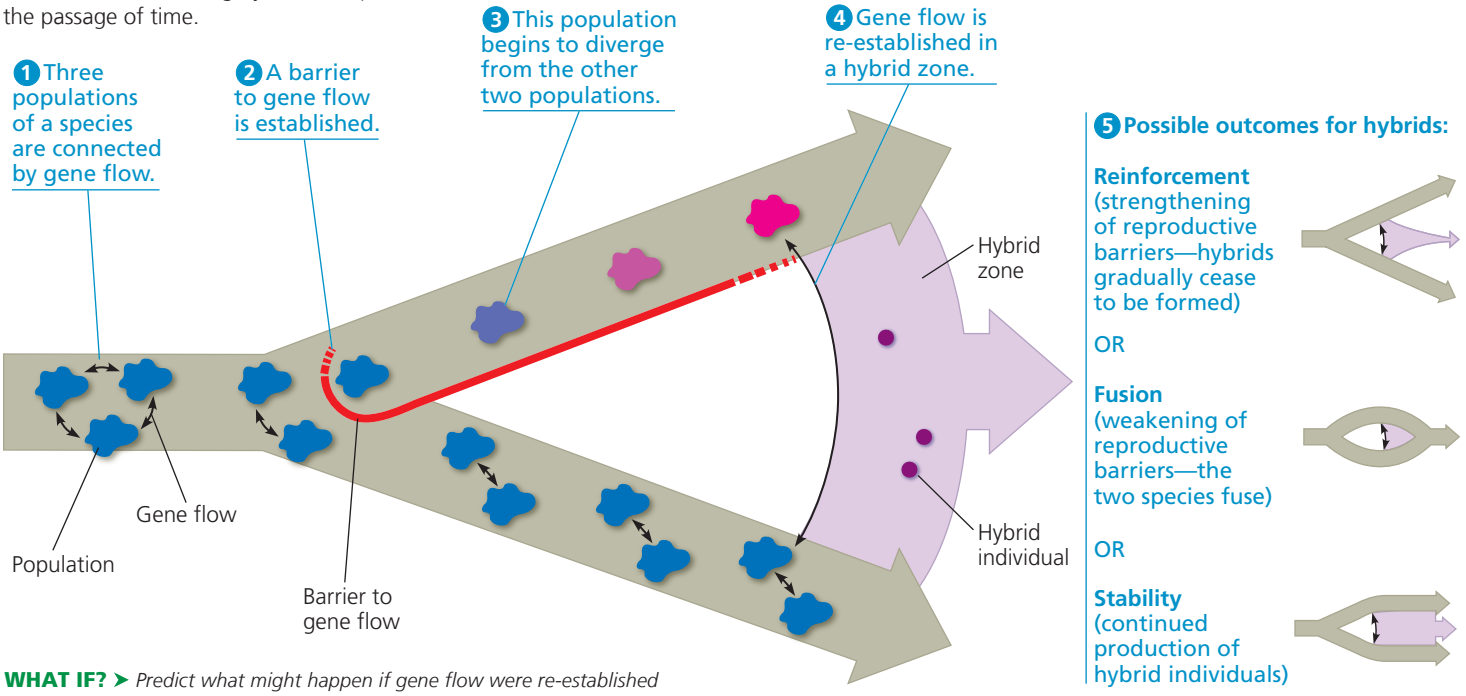
barriers, fusion of species, or stability **(Figure 24.14)**. Let's examine what studies suggest about these possibilities.

Reinforcement: Strengthening Reproductive Barriers

Hybrids often are less fit than members of their parent species. In such cases, natural selection should strengthen prezygotic barriers to reproduction, reducing the formation of unfit hybrids. Because this process involves *reinforcing* reproductive barriers, it is called **reinforcement**. If reinforcement is occurring, a logical prediction is that barriers to reproduction between species should be stronger for sympatric populations than for allopatric populations.

As an example, let's consider two species of European flycatcher, the pied flycatcher (*Ficedula hypoleuca*) and the collared flycatcher (*Ficedula albicollis*). In allopatric populations of these birds, males of the two species closely resemble one another, while in sympatric populations, the males look very different. Female flycatchers do not select males of the other species when given a choice between males from sympatric populations, but they frequently do make mistakes when selecting between males from allopatric populations. Thus, barriers to reproduction are stronger in birds from sympatric populations than in birds from allopatric populations, as you would predict if reinforcement were occurring. Similar results have been observed in a number of organisms, including fishes, insects, plants, and other birds.

Figure 24.14 Formation of a hybrid zone and possible outcomes for hybrids over time. The thick gray arrows represent the passage of time.



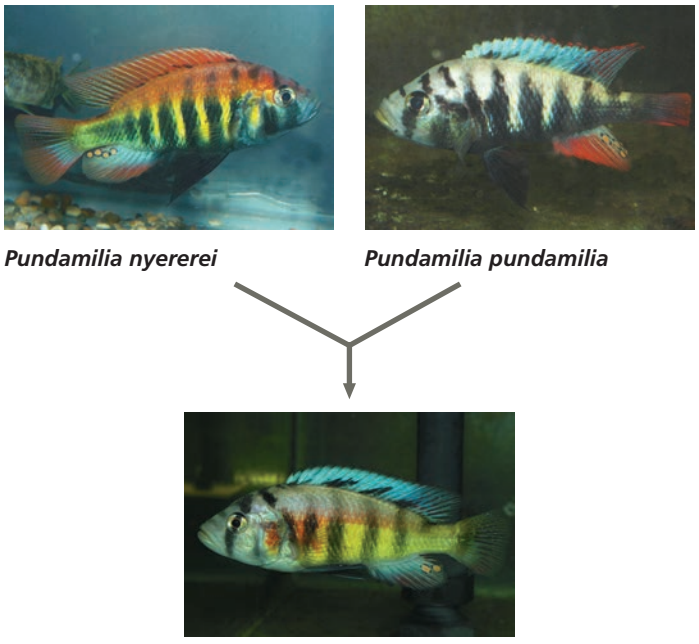
WHAT IF? ➤ *Predict what might happen if gene flow were re-established at step 3 in this process.*

Fusion: Weakening Reproductive Barriers

Barriers to reproduction may be weak when two species meet in a hybrid zone. Indeed, so much gene flow may occur that reproductive barriers weaken further and the gene pools of the two species become increasingly alike. In effect, the speciation process reverses, eventually causing the two hybridizing species to fuse into a single species.

For example, genetic and morphological evidence indicate that the recent loss of the large tree finch from the Galápagos island of Floreana resulted from extensive hybridization with another finch species on that island. Such a situation also may be occurring among Lake Victoria cichlids. Many pairs of ecologically similar cichlid species are reproductively isolated because the females of one species prefer to mate with males of one color, while females of the other species prefer to mate with males of a different color (see Figure 24.12). Results from field and laboratory studies indicate that murky waters caused by pollution have reduced the ability of females to use color to distinguish males of their own species from males of closely related species. In some polluted waters, many hybrids have been produced, leading to fusion of the parent species' gene pools and a loss of species **(Figure 24.15)**.

Figure 24.15 Fusion: the breakdown of reproductive barriers. Increasingly cloudy water in Lake Victoria over the past several decades may have weakened reproductive barriers between *P. nyererei* and *P. pundamilia*. In areas of cloudy water, the two species have hybridized extensively, causing their gene pools to fuse.



Pundamilia nyererei ***Pundamilia pundamilia***

Pundamilia "turbid water," hybrid offspring from a location with turbid water

Stability: Continued Formation of Hybrid Individuals

Many hybrid zones are stable in the sense that hybrids continue to be produced. In some cases, this occurs because the hybrids survive or reproduce better than members of either parent species, at least in certain habitats or years. But stable hybrid zones have also been observed in cases where the hybrids are selected *against*—an unexpected result.

For example, hybrids continue to form in the *Bombina* hybrid zone even though they are strongly selected against.

One explanation relates to the narrowness of the *Bombina* hybrid zone (see Figure 24.13). Evidence suggests that members of both parent species migrate into the zone from the parent populations located outside the zone, thus leading to the continued production of hybrids. If the hybrid zone were wider, this would be less likely to occur, since the center of the zone would receive little gene flow from distant parent populations located outside the hybrid zone.

Sometimes the outcomes in hybrid zones match our predictions (European flycatchers and cichlid fishes), and sometimes they don't (*Bombina*). But whether our predictions are upheld or not, events in hybrid zones can shed light on how barriers to reproduction between closely related species change over time. In the next section, we'll examine how interactions between hybridizing species can also provide a glimpse into the speed and genetic control of speciation.

CONCEPT CHECK 24.3

1. What are hybrid zones, and why can they be viewed as "natural laboratories" in which to study speciation?
2. **WHAT IF?** ➤ Consider two species that diverged while geographically separated but resumed contact before reproductive isolation was complete. Predict the outcome over time if the two species mated indiscriminately and (a) hybrid offspring survived and reproduced more poorly than offspring from intraspecific matings or (b) hybrid offspring survived and reproduced as well as offspring from intraspecific matings.

For suggested answers, see Appendix A.

CONCEPT 24.4

Speciation can occur rapidly or slowly and can result from changes in few or many genes

Darwin faced many questions when he began to ponder that "mystery of mysteries"—speciation. He found answers to some of those questions when he realized that evolution by natural selection helps explain both the diversity of life and the adaptations of organisms (see Concept 21.2). But biologists since Darwin have continued to ask fundamental questions about speciation. How long does it take for new species to form? And how many genes change when one species splits into two? Answers to these questions are also emerging.

The Time Course of Speciation

We can gather information about how long it takes new species to form from broad patterns in the fossil record and from studies that use morphological data (including fossils) or molecular data to assess the time interval between speciation events in particular groups of organisms.

Patterns in the Fossil Record

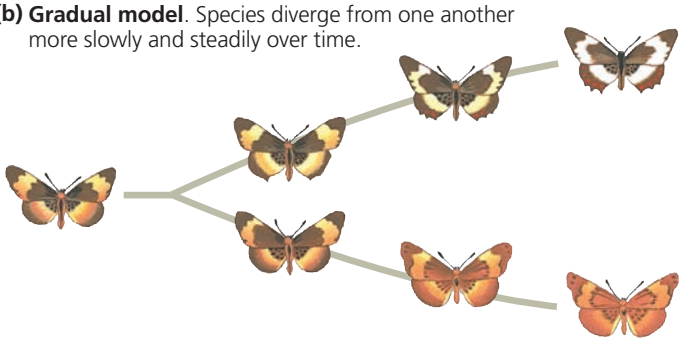
The fossil record includes many episodes in which new species appear suddenly in a geologic stratum, persist essentially unchanged through several strata, and then disappear. For example, there are dozens of species of marine invertebrates that make their debut in the fossil record with novel morphologies, but then change little for millions of years before becoming extinct. The term **punctuated equilibria** is used to describe these periods of apparent stasis punctuated by sudden change **(Figure 24.16a)**. Other species do not show a punctuated pattern; instead, they appear to have changed more gradually over long periods of time **(Figure 24.16b)**.

What might punctuated and gradual patterns tell us about how long it takes new species to form? Suppose that a species survived for 5 million years, but most of the morphological changes that caused it to be designated a new species occurred during the first 50,000 years of its existence—just 1% of its total lifetime. Time periods this short (in geologic terms) often cannot be distinguished in fossil strata, in part because the rate of sediment accumulation may be too slow to separate layers this close in time. Thus, based on its fossils, the species would seem to have appeared suddenly and then lingered with little or no change before becoming extinct. Even though such a species may have originated more slowly than its fossils suggest (in this case taking up to 50,000 years), a punctuated pattern indicates that speciation occurred relatively rapidly. For species whose fossils changed much more gradually, we also cannot tell exactly when a new biological species formed, since information about reproductive isolation does not fossilize.

Figure 24.16 Two models for the tempo of speciation.

(a) Punctuated model. New species change most as they branch from a parent species and then change little for the rest of their existence.

(b) Gradual model. Species diverge from one another more slowly and steadily over time.



However, it is likely that speciation in such groups occurred relatively slowly, perhaps taking millions of years.

Interview with Stephen Jay Gould: An "architect" of the concept of punctuated equilibria

Speciation Rates

The existence of fossils that display a punctuated pattern suggests that once the process of speciation begins, it can be completed relatively rapidly—a suggestion supported by a growing number of studies.

For example, rapid speciation appears to have produced the wild sunflower *Helianthus anomalus*. Genetic evidence indicates that this species originated by the hybridization of two other sunflower species, *H. annuus* and *H. petiolaris*. The hybrid species *H. anomalus* is ecologically distinct and reproductively isolated from both parent species **(Figure 24.17)**. Unlike the outcome of allopolyploid speciation, in which there is a change in chromosome number after hybridization, in these sunflowers the two parent species and the hybrid all have the same number of chromosomes ($2n = 34$). How, then, did speciation occur? To study this question, researchers performed an experiment designed to mimic events in nature **(Figure 24.18)**. Their results indicated that natural selection could produce extensive genetic changes in hybrid populations over short periods of time. These changes appear to have caused the hybrids to diverge reproductively from their parents and form a new species, *H. anomalus*.

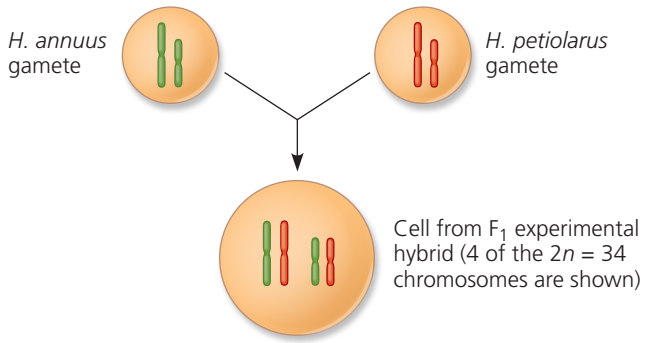
The sunflower example, along with the apple maggot fly, Lake Victoria cichlid, and fruit fly examples discussed earlier, suggests that new species can arise rapidly *once divergence begins*. But what is the total length of time between speciation events? This interval consists of the time that elapses before populations of a newly formed species start to diverge from one another plus the time it takes for speciation to be complete once divergence begins. It turns out that the total time between speciation events varies considerably. In a survey of data from 84 groups of plants and animals, speciation intervals ranged from 4,000 years (in cichlids of Lake Nabugabo,

Figure 24.17 A hybrid sunflower species and its dry sand dune habitat. The wild sunflower *Helianthus anomalus* shown here originated via the hybridization of two other sunflowers, *H. annuus* and *H. petiolaris*, which live in nearby but moister environments.

Figure 24.18

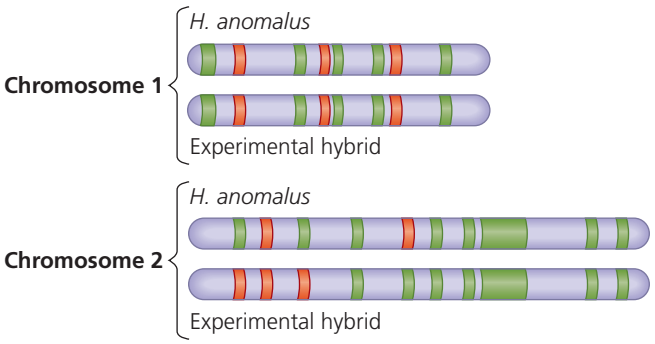
Inquiry How does hybridization lead to speciation in sunflowers?

Experiment Loren Rieseberg and his colleagues crossed the two parent sunflower species, *H. annuus* and *H. petiolaris*, to produce experimental hybrids in the laboratory (for each gamete, only two of the $n = 17$ chromosomes are shown).



Note that in the first (F_1) generation, each chromosome of the experimental hybrids consisted entirely of DNA from one or the other parent species. The researchers then tested whether the F_1 and subsequent generations of experimental hybrids were fertile. They also used species-specific genetic markers to compare the chromosomes in the experimental hybrids with the chromosomes in the naturally occurring hybrid *H. anomalus*.

Results Although only 5% of the F_1 experimental hybrids were fertile, after just four more generations the hybrid fertility rose to more than 90%. The chromosomes of individuals from this fifth hybrid generation differed from those in the F_1 generation (see above) but were similar to those in *H. anomalus* individuals from natural populations:



Conclusion Over time, the chromosomes in the population of experimental hybrids became similar to the chromosomes of *H. anomalus* individuals from natural populations. This suggests that the observed rise in the fertility of the experimental hybrids may have occurred as selection eliminated regions of DNA from the parent species that were not compatible with one another. Overall, it appeared that the initial steps of the speciation process occurred rapidly and could be mimicked in a laboratory experiment.

Data from L. H. Rieseberg et al., Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids, *Science* 272:741–745 (1996).

WHAT IF? *The increased fertility of the experimental hybrids could have resulted from natural selection for thriving under laboratory conditions. Evaluate this alternative explanation for the result.*

Uganda) to 40 million years (in some beetles). Overall, the time between speciation events averaged 6.5 million years and was rarely less than 500,000 years.

These data suggest that on average, millions of years may pass before a newly formed plant or animal species will itself give rise to another new species. As you'll read in Concept 25.4, this finding has implications for how long it takes life on Earth to recover from mass extinction events. Moreover, the extreme variability in the time it takes new species to form indicates that organisms do not have an internal "speciation clock" that causes them to produce new species at regular intervals. Instead, speciation begins only after gene flow between populations is interrupted, perhaps by changing environmental conditions or by unpredictable events, such as a storm that transports a few individuals to a new area. Furthermore, once gene flow is interrupted, the populations must diverge genetically to such an extent that they become reproductively isolated—all before other events cause gene flow to resume, possibly reversing the speciation process (see Figure 24.15).

Studying the Genetics of Speciation

Studies of ongoing speciation (as in hybrid zones) can reveal traits that cause reproductive isolation. By identifying the genes that control those traits, scientists can explore a fundamental question of evolutionary biology: How many genes influence the formation of new species?

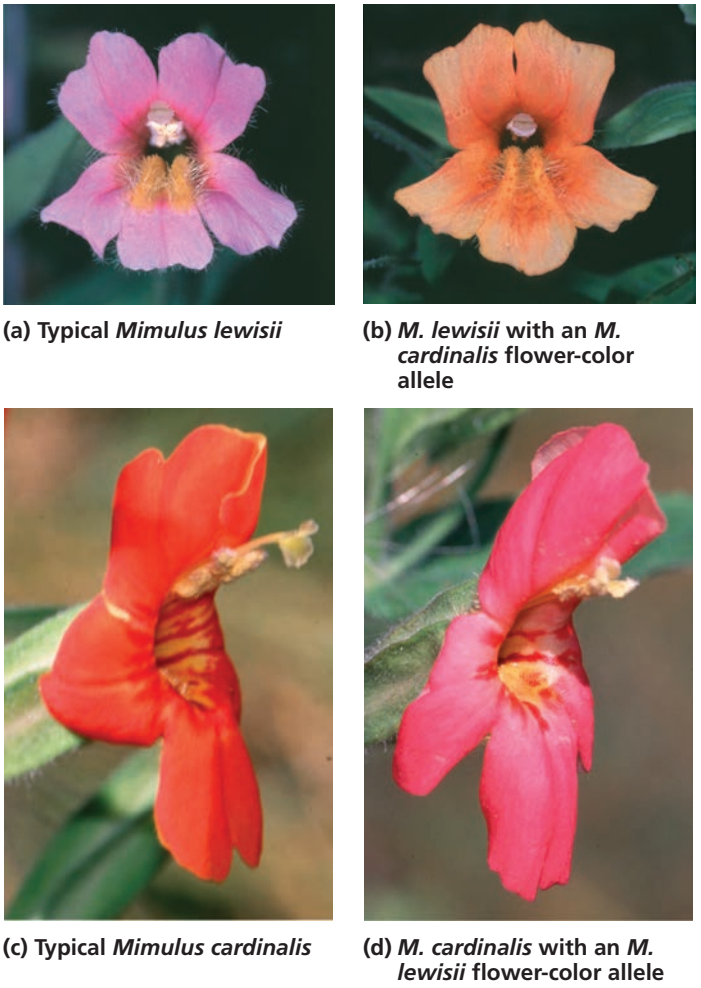
In some cases, the evolution of reproductive isolation results from the effects of a single gene. For example, in Japanese snails of the genus *Euhadra*, a change in a single gene results in a mechanical barrier to reproduction. This gene controls the direction in which the shells spiral. When their shells spiral in different directions, the snails' genitalia are oriented in a manner that prevents mating (Figure 24.3f shows a similar example). Recent genetic analyses have uncovered other single genes that cause reproductive isolation in fruit flies or mice.

A major barrier to reproduction between two closely related species of monkey flower, *Mimulus cardinalis* and *M. lewisii*, also appears to be influenced by a relatively small number of genes. These two species are isolated by several prezygotic and postzygotic barriers. Of these, one prezygotic barrier, pollinator choice, accounts for most of the isolation: In a hybrid zone between *M. cardinalis* and *M. lewisii*, nearly 98% of pollinator visits were restricted to one species or the other.

The two monkey flower species are visited by different pollinators: Hummingbirds prefer the red-flowered *M. cardinalis*, and bumblebees prefer the pink-flowered *M. lewisii*. Pollinator choice is affected by at least two loci in the monkey flowers, one of which, the "yellow upper," or *yup*, locus, influences flower color **(Figure 24.19)**. By crossing the two parent species to produce F_1 hybrids and then performing repeated backcrosses of these F_1 hybrids to each parent species, researchers succeeded in transferring the *M. cardinalis* allele at this locus into *M. lewisii*, and vice versa. In a field experiment, *M. lewisii* plants with the *M. cardinalis yup* allele received 68-fold more visits from hummingbirds than did wild-type *M. lewisii*. Similarly, *M. cardinalis* plants with the *M. lewisii yup* allele received 74-fold more visits from bumblebees than did wild-type *M. cardinalis*. Thus, a mutation at a single locus can influence pollinator preference and hence contribute to reproductive isolation in monkey flowers.

In other organisms, the speciation process is influenced by larger numbers of genes and gene interactions. For example, hybrid sterility between two subspecies of the fruit fly *Drosophila pseudoobscura* results from gene interactions among at least four loci, and postzygotic isolation in the sunflower hybrid zone discussed earlier is influenced by at least 26 chromosome segments (and an unknown number of genes). Overall, studies suggest that few or many genes can influence the evolution of reproductive isolation and hence the emergence of a new species.

Figure 24.19 A locus that influences pollinator choice. Pollinator preferences provide a strong barrier to reproduction between *Mimulus lewisii* and *M. cardinalis*. After transferring the *M. lewisii* allele for a flower-color locus into *M. cardinalis* and vice versa, researchers observed a shift in some pollinators' preferences.



(a) Typical *Mimulus lewisii*

(b) *M. lewisii* with an *M. cardinalis* flower-color allele

(c) Typical *Mimulus cardinalis*

(d) *M. cardinalis* with an *M. lewisii* flower-color allele

WHAT IF ➤ *If* M. cardinalis *individuals that had the* M. lewisii yup *allele were planted in an area that housed both monkey flower species, how might the production of hybrid offspring be affected?*

From Speciation to Macroevolution

As you've seen, speciation may begin with differences as small as the color on a cichlid's back. However, as speciation occurs again and again, such differences can accumulate and become more pronounced, eventually leading to the formation of new groups of organisms that differ greatly from their ancestors (as in the origin of whales from terrestrial mammals; see Figure 21.20). Moreover, as one group of organisms increases in size by producing many new species, another group of organisms may shrink, losing species to extinction. The cumulative effects of many such speciation and extinction events have helped shape the sweeping evolutionary changes that are documented in the fossil record. In the next chapter, we turn to such large-scale evolutionary changes as we begin our study of macroevolution.

CONCEPT CHECK 24.4

1. Speciation can occur rapidly between diverging populations, yet the time between speciation events is often more than a million years. Explain this apparent contradiction.
2. Summarize evidence that the *yup* locus acts as a prezygotic barrier to reproduction in two species of monkey flowers. Do these results demonstrate that the *yup* locus alone controls barriers to reproduction between these species? Explain.
3. **MAKE CONNECTIONS ➤** Compare Figure 13.12 with Figure 24.18. What cellular process could cause the hybrid chromosomes in Figure 24.18 to contain DNA from both parent species? Explain.

For suggested answers, see Appendix A.

24 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 24.1

The biological species concept emphasizes reproductive isolation *(pp. 561–564)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

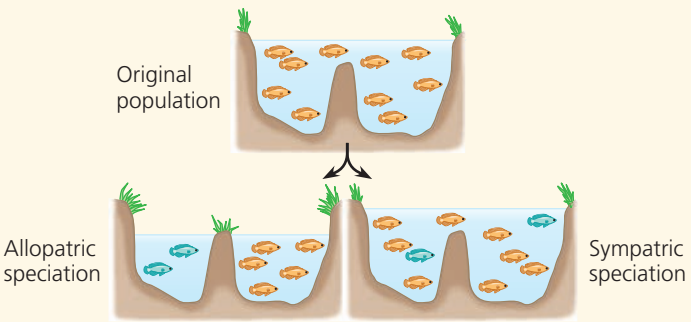
- A biological **species** is a group of populations whose individuals may interbreed and produce viable, fertile offspring with each other but not with members of other species.
- The **biological species concept** emphasizes reproductive isolation through **prezygotic** and **postzygotic barriers** that separate gene pools.

? *Explain the role of gene flow in the biological species concept.*

CONCEPT 24.2

Speciation can take place with or without geographic separation *(pp. 565–570)*

- In **allopatric speciation**, gene flow is reduced when two populations of one species become geographically separated from each other. One or both populations may undergo evolutionary change during the period of separation, resulting in the establishment of barriers to reproduction.



- In **sympatric speciation**, a new species originates while remaining in the same geographic area as the parent species. Plant species (and, more rarely, animal species) have evolved sympatrically through **polyploidy**. Sympatric speciation can also result from sexual selection and habitat shifts.

? *Can factors that cause sympatric speciation also cause allopatric speciation? Explain.*

CONCEPT 24.3

Hybrid zones reveal factors that cause reproductive isolation *(pp. 570–574)*

- Many groups of organisms form **hybrid zones** in which members of different species meet and mate, producing at least some offspring of mixed ancestry.
- Many hybrid zones are *stable,* in that hybrid offspring continue to be produced over time. In others, **reinforcement** strengthens prezygotic barriers to reproduction, thus decreasing the formation of unfit hybrids. In still other hybrid zones, barriers to reproduction may weaken over time, resulting in the *fusion* of the species' gene pools (reversing the speciation process).

? *What factors can support the long-term stability of a hybrid zone if the parent species live in different environments?*

CONCEPT 24.4

Speciation can occur rapidly or slowly and can result from changes in few or many genes *(pp. 574–577)*

- New species can form rapidly once divergence begins—but it can take millions of years for that to happen. The time interval between speciation events varies considerably, from a few thousand years to tens of millions of years.
- Researchers have identified particular genes involved in some cases of speciation. Speciation can be driven by few or many genes.

? *Is speciation something that happened only in the distant past, or are new species continuing to arise today? Explain.*

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–6 can be found in the Study Area in MasteringBiology.

PRACTICE TEST goo.gl/iAsVgL

7. **EVOLUTION CONNECTION** Explain the biological basis for assigning all human populations to a single species. Can you think of a scenario by which a second human species could originate in the future?

8. **SCIENTIFIC INQUIRY • DRAW IT** In this chapter, you read that bread wheat (*Triticum aestivum*) is an allohexaploid, containing two sets of chromosomes from each of three different parent species. Genetic analysis suggests that the three species pictured following this question each contributed chromosome sets to *T. aestivum*. (The capital letters here represent sets of chromosomes rather than individual genes, and the diploid chromosome number for each species is shown in parentheses.) Evidence also indicates that the first polyploidy event was a spontaneous hybridization of the early cultivated wheat species *T. monococcum* and a wild *Triticum* grass species. Based on this information, draw a diagram of one possible chain of events that could have produced the allohexaploid *T. aestivum*.

Ancestral species:

AA

Triticum monococcum (14)

BB

Wild *Triticum* (14)

DD

Wild *T. tauschii* (14)

Product:

AA BB DD

T. aestivum (bread wheat) (42)

9. **WRITE ABOUT A THEME: INFORMATION** In sexually reproducing species, each individual inherits DNA from both parent organisms. In a short essay (100–150 words), apply this idea to what occurs when organisms of two species that have homologous chromosomes mate and produce (F_1) hybrid offspring. What percentage of the DNA in the F_1 hybrids' chromosomes comes from each parent species? As the hybrids mate and produce F_2 and later-generation hybrid offspring, describe how recombination and natural selection may affect whether the DNA in hybrid chromosomes is derived from one parent species or the other.

10. **SYNTHESIZE YOUR KNOWLEDGE**

Suppose that females of one population of strawberry poison dart frogs (*Dendrobates pumilio*) prefer to mate with males that are orange-red in color. In a different population, females prefer males with yellow skin. Explain how such differences could arise and how they could affect the evolution of reproductive isolation in allopatric versus sympatric populations.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

25 Macroevolution

Figure 25.1 Would you have expected to find whale bones buried here?

KEY CONCEPTS

Fossil of *Dorudon atrox*, an ancient whale

A Surprise in the Desert

With its dry, wind-sculpted sands and searing heat, the Sahara Desert seems an unlikely place to discover the bones of whales. But starting in the 1870s, researchers uncovered fossils of ancient whales at several locations that once were covered by an ancient sea **(Figure 25.1)**. For example, a nearly complete skeleton of *Dorudon atrox*, an extinct whale that lived 35 million years ago, was discovered in a region that came to be called Wadi Hitan, the "Valley of Whales." Collectively, the whale fossils found in the Sahara were spectacular not only for where they were found, but also for documenting early steps in the transition from life on land to life in the sea.

Fossils discovered in other parts of the world tell a similar story: Past organisms were very different from those presently living. The sweeping changes in life on Earth as revealed by fossils illustrate **macroevolution**, the broad pattern of evolution above the species level. Examples of macroevolutionary change include the emergence of terrestrial vertebrates through a series of speciation events, the impact of mass extinctions on biodiversity, and the origin of key adaptations such as flight.

Taken together, such changes provide a grand view of the evolutionary history of life. We'll examine that history in this chapter, beginning with a discussion of hypotheses regarding the origin of life. This is the most speculative topic of the entire unit, for no fossil evidence of that seminal episode exists. We will then turn to evidence from the fossil record about major events in the history of life and the factors that have shaped the rise and fall of different groups of organisms over time.

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

CONCEPT 25.1

Conditions on early Earth made the origin of life possible

Direct evidence of life on early Earth comes from fossils of microorganisms that lived 3.5 billion years ago. But how did the first living cells appear? Observations and experiments in chemistry, geology, and physics have led scientists to propose one scenario that we'll examine here. They hypothesize that chemical and physical processes could have produced simple cells through a sequence of four main stages:

1. The abiotic (nonliving) synthesis of small organic molecules, such as amino acids and nitrogenous bases
2. The joining of these small molecules into macromolecules, such as proteins and nucleic acids
3. The packaging of these molecules into **protocells**, droplets with membranes that maintained an internal chemistry different from that of their surroundings
4. The origin of self-replicating molecules that eventually made inheritance possible

Though speculative, this scenario leads to predictions that can be tested in the laboratory. In this section, we'll examine some of the evidence for each stage.

Synthesis of Organic Compounds on Early Earth

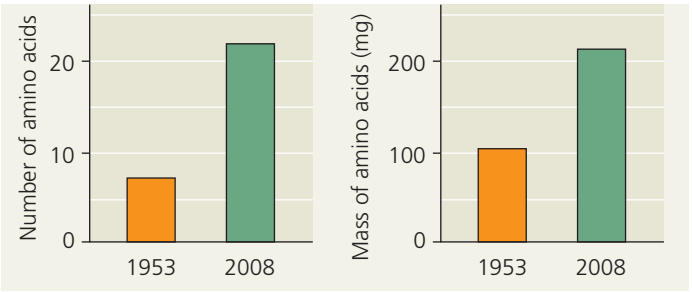
Our planet formed 4.6 billion years ago, condensing from a vast cloud of dust and rocks that surrounded the young sun. For its first few hundred million years, Earth was bombarded by huge chunks of rock and ice left over from the formation of the solar system. The collisions generated so much heat that all of the available water was vaporized, preventing the formation of seas and lakes.

This massive bombardment ended 4 billion years ago, setting the stage for the origin of life. The first atmosphere had little oxygen and was likely thick with water vapor, along with compounds released by volcanic eruptions, such as nitrogen and its oxides, carbon dioxide, methane, ammonia, and hydrogen. As Earth cooled, the water vapor condensed into oceans, and much of the hydrogen escaped into space.

During the 1920s, Russian chemist A. I. Oparin and British scientist J. B. S. Haldane independently hypothesized that Earth's early atmosphere was a reducing (electron-adding) environment, in which organic compounds could have formed from simpler molecules. The energy for this synthesis could have come from lightning and UV radiation. Haldane suggested that the early oceans were a solution of organic molecules, a "primitive soup" from which life arose.

In 1953, Stanley Miller and Harold Urey, working at the University of Chicago, tested the Oparin-Haldane hypothesis by creating laboratory conditions comparable to those that scientists at the time thought existed on early Earth (see Figure 4.2). His apparatus yielded a variety of amino acids found in organisms today, along with other organic compounds. Many laboratories have since repeated Miller's classic experiment using different recipes for the atmosphere, some of which also produced organic compounds.

Figure 25.2 Amino acid synthesis in a simulated volcanic eruption. In addition to his classic 1953 study, Miller also conducted an experiment simulating a volcanic eruption. In a 2008 reanalysis of those results, researchers found that far more amino acids were produced under simulated volcanic conditions than were produced in the conditions of the original 1953 experiment.



MAKE CONNECTIONS *Explain how more than 20 amino acids could have been produced in the 2008 experiment. (See Concept 5.4.)*

However, some evidence suggests that the early atmosphere was made up primarily of nitrogen and carbon dioxide and was neither reducing nor oxidizing (electron removing). Recent Miller/Urey-type experiments using such "neutral" atmospheres have also produced organic molecules. In addition, small pockets of the early atmosphere, such as those near the openings of volcanoes, may have been reducing. Perhaps the first organic compounds formed near volcanoes. In 2008, researchers used modern equipment to reanalyze molecules that Miller had saved from one of his experiments. The 2008 paper found that numerous amino acids had formed under conditions that simulated a volcanic eruption **(Figure 25.2)**.

Another hypothesis is that organic compounds were first produced in deep-sea **hydrothermal vents**, areas on the seafloor where heated water and minerals gush from Earth's interior into the ocean. Some of these vents, known as "black smokers," release water so hot (300–400°C) that organic compounds formed there may have been unstable. But other deep-sea vents, called **alkaline vents**, release water that has a high pH (9–11) and is warm (40–90°C) rather than hot, an environment that may have been more suitable for the origin of life **(Figure 25.3)**.

Studies related to the volcanic-atmosphere and alkaline-vent hypotheses show that the abiotic synthesis of organic molecules is possible under various conditions. Another source of organic molecules may have been meteorites. For example, fragments of the Murchison meteorite, a 4.5-billion-year-old rock that landed in Australia in 1969, contain more than 80 amino acids, some in large amounts. These amino

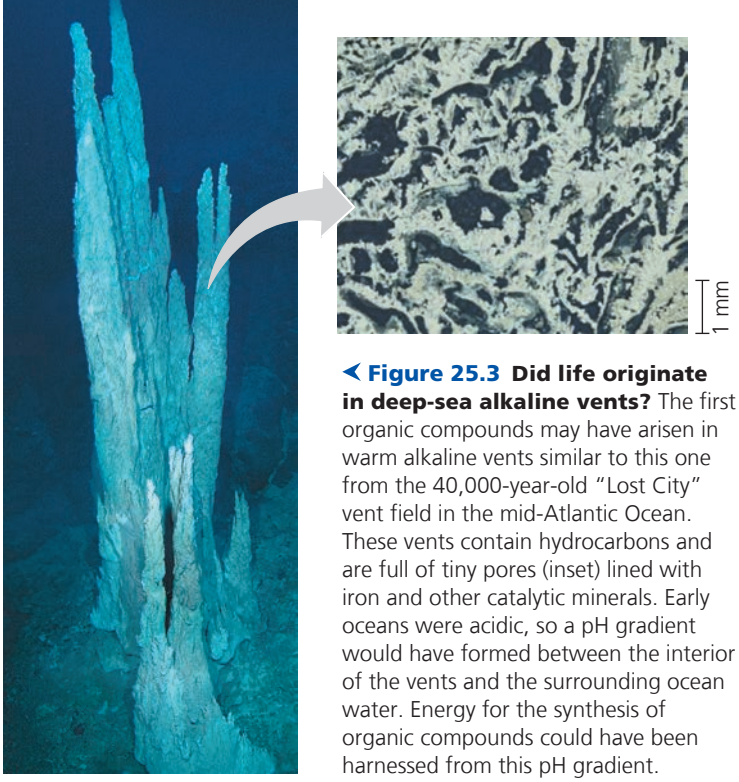


Figure 25.3 Did life originate in deep-sea alkaline vents? The first organic compounds may have arisen in warm alkaline vents similar to this one from the 40,000-year-old "Lost City" vent field in the mid-Atlantic Ocean. These vents contain hydrocarbons and are full of tiny pores (inset) lined with iron and other catalytic minerals. Early oceans were acidic, so a pH gradient would have formed between the interior of the vents and the surrounding ocean water. Energy for the synthesis of organic compounds could have been harnessed from this pH gradient.

acids cannot be contaminants from Earth because they consist of an equal mix of D and L isomers (see Figure 4.7). Organisms make and use only L isomers, with a few rare exceptions. Recent studies have shown that the Murchison meteorite also contained other key organic molecules, including lipids, simple sugars, and nitrogenous bases such as uracil.

Abiotic Synthesis of Macromolecules

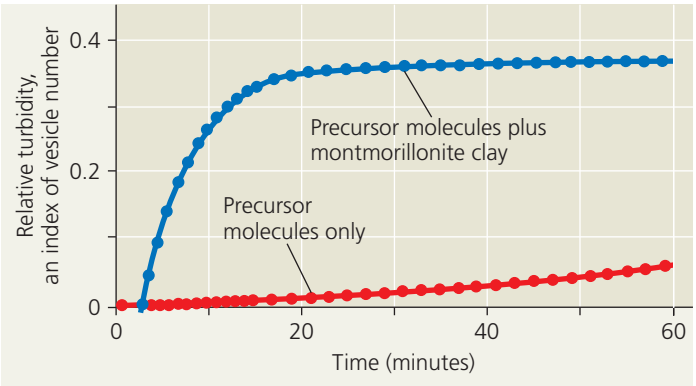
The presence of small organic molecules, such as amino acids and nitrogenous bases, is not sufficient for the emergence of life as we know it. Every cell has many types of macromolecules, including enzymes and other proteins and the nucleic acids needed for self-replication. Could such macromolecules have formed on early Earth? A 2009 study demonstrated that one key step, the abiotic synthesis of RNA monomers, can occur spontaneously from simple precursor molecules. In addition, by dripping solutions of amino acids or RNA nucleotides onto hot sand, clay, or rock, researchers have produced polymers of these molecules. The polymers formed spontaneously, without the help of enzymes or ribosomes. Unlike proteins, the amino acid polymers are a complex mix of linked and cross-linked amino acids. Still, it is possible that such polymers acted as weak catalysts for a variety of chemical reactions on early Earth.

Protocells

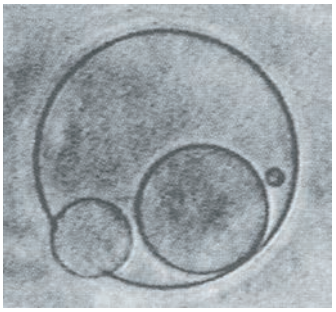
All organisms must be able to carry out both reproduction and energy processing (metabolism). DNA molecules carry genetic information, including the instructions needed to replicate themselves accurately during reproduction. But DNA replication requires elaborate enzymatic machinery, along with an abundant supply of nucleotide building blocks provided by the cell's metabolism. This suggests that self-replicating molecules and a metabolic source of building blocks may have appeared together in early protocells. The necessary conditions may have been met in *vesicles*, fluid-filled compartments enclosed by a membrane-like structure. Recent experiments show that abiotically produced vesicles can exhibit certain properties of life, including simple reproduction and metabolism, as well as the maintenance of an internal chemical environment different from that of their surroundings **(Figure 25.4)**.

For example, vesicles can form spontaneously when lipids or other organic molecules are added to water. When this occurs, molecules that have both a hydrophobic region and a hydrophilic region can organize into a bilayer similar to the lipid bilayer of a plasma membrane. Adding substances such as *montmorillonite*, a soft mineral clay produced by the weathering of volcanic ash, greatly increases the rate of vesicle self-assembly (see Figure 25.4a). This clay, which is thought to have been common on early Earth, provides surfaces on which organic molecules become concentrated, increasing the likelihood that the molecules will react with each other and form vesicles. Abiotically produced vesicles can "reproduce" on their own (see Figure 25.4b), and they can increase

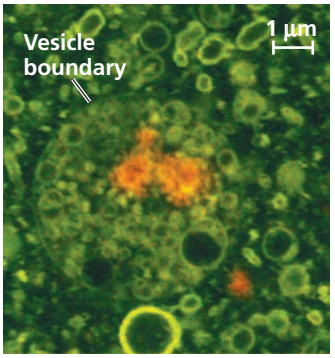
Figure 25.4 Features of abiotically produced vesicles.



(a) Self-assembly. The presence of montmorillonite clay greatly increases the rate of vesicle self-assembly.



(b) Reproduction. Vesicles can divide on their own, as in this vesicle "giving birth" to smaller vesicles (LM).



(c) Absorption of RNA. This vesicle has incorporated montmorillonite clay particles coated with RNA (orange).

MAKE CONNECTIONS *Explain how molecules with both a hydrophobic region and a hydrophilic region can self-assemble into a bilayer when in water. (See Concept 5.3.)*

in size ("grow") without dilution of their contents. Vesicles also can absorb montmorillonite particles, including those on which RNA and other organic molecules have become attached (see Figure 25.4c). Finally, experiments have shown that some vesicles have a selectively permeable bilayer and can perform metabolic reactions using an external source of reagents—another important prerequisite for life.

Self-Replicating RNA

The first genetic material was most likely RNA, not DNA. RNA plays a central role in protein synthesis, but it can also function as an enzyme-like catalyst (see Concept 17.3). Such RNA catalysts are called **ribozymes**. Some ribozymes can make complementary copies of short pieces of RNA, provided that they are supplied with nucleotide building blocks.

Natural selection on the molecular level has produced ribozymes capable of self-replication in the laboratory. How does this occur? Unlike double-stranded DNA, which takes the form of a uniform helix, single-stranded RNA molecules assume a variety of specific three-dimensional shapes mandated by their nucleotide sequences. In a given environment, RNA molecules with certain nucleotide sequences may have shapes that enable them to replicate faster and with fewer errors than other sequences. The RNA molecule with the greatest ability to replicate itself will leave the most descendant molecules. Occasionally, a copying error will result in a molecule with a shape that is even more adept at self-replication. Similar selection events may have occurred on early Earth. Thus, life as we know it may have been preceded by an "RNA world," in which small RNA molecules were able to replicate and to store genetic information about the vesicles that carried them.

In 2013, Dr. Jack Szostak and colleagues succeeded in building a vesicle in which copying of a template strand of RNA could occur—a key step towards constructing a vesicle with self-replicating RNA. On early Earth, a vesicle with such self-replicating, catalytic RNA would differ from its many neighbors that lacked such molecules. If that vesicle could grow, split, and pass its RNA molecules to its "daughters," the daughters would be protocells. Although the first such protocells likely carried only limited amounts of genetic information, specifying only a few properties, their inherited characteristics could have been acted on by natural selection. The most successful of the early protocells would have increased in number because they could exploit their resources effectively and pass their abilities on to subsequent generations.

Once RNA sequences that carried genetic information appeared in protocells, many additional changes would have been possible. For example, RNA could have provided the template on which DNA nucleotides were assembled. Double-stranded DNA is a more chemically stable repository for genetic information than is the more fragile RNA. DNA also can be replicated more accurately. Accurate replication was advantageous as genomes grew larger through gene duplication and other processes and as more properties of the protocells became coded in genetic information. Once DNA appeared, the stage was set for a blossoming of new forms of life—a change we see documented in the fossil record.

Interview with Jack Szostak: Studying the origin of life

CONCEPT CHECK 25.1

1. What conditions on early Earth could have permitted the synthesis of organic molecules?
2. How would the appearance of protocells have represented a key step in the origin of life?
3. **MAKE CONNECTIONS ➤** In changing from an "RNA world" to today's "DNA world," genetic information must have flowed from RNA to DNA. After reviewing Figures 17.4 and 26.9, suggest how this could have occurred. Does such a flow occur today?

For suggested answers, see Appendix A.

CONCEPT 25.2

The fossil record documents the history of life

Starting with the earliest traces of life, the fossil record opens a window into the world of long ago and provides glimpses of the evolution of life over billions of years. In this section, we'll examine fossils as a form of scientific evidence: how fossils form, how scientists date and interpret them, and what they can and cannot tell us about changes in the history of life.

The Fossil Record

Sedimentary rocks are the richest source of fossils. As a result, the fossil record is based primarily on the sequence in which fossils have accumulated in sedimentary rock layers, called *strata* (see Figure 21.3). Useful information is also provided by other types of fossils, such as insects preserved in amber (fossilized tree sap) and mammals frozen in ice.

The fossil record shows that there have been great changes in the kinds of organisms on Earth at different points in time **(Figure 25.5)**. Many past organisms were unlike organisms living today, and many organisms that once were common are now extinct. As we'll see later in this section, fossils also document how new groups of organisms arose from previously existing ones.

As substantial and significant as the fossil record is, keep in mind that it is an incomplete chronicle of evolution. Many of Earth's organisms did not die in the right place and time to be preserved as fossils. Of those fossils that were formed, many were destroyed by later geologic processes, and only a fraction of the others have been discovered. As a result, the known fossil record is biased in favor of species that existed for a long time, were abundant and widespread in certain kinds of environments, and had hard shells, skeletons, or other parts that

Figure 25.5 Documenting the history of life. These fossils illustrate representative organisms from different points in time. Although prokaryotes and unicellular eukaryotes are shown only at the base of the diagram, these organisms continue to thrive today. In fact, most organisms on Earth are unicellular.

Present

100 million years ago

175

200

270

300

375

400

500

510

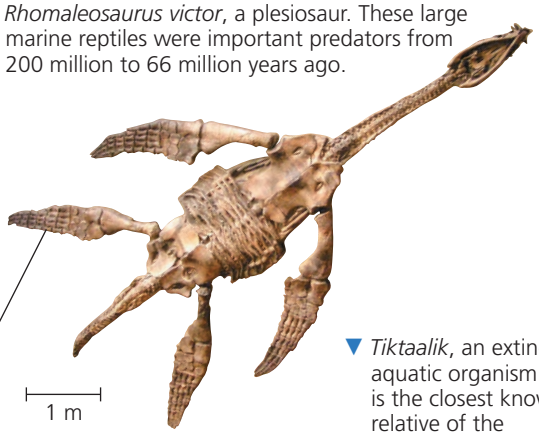
560

600

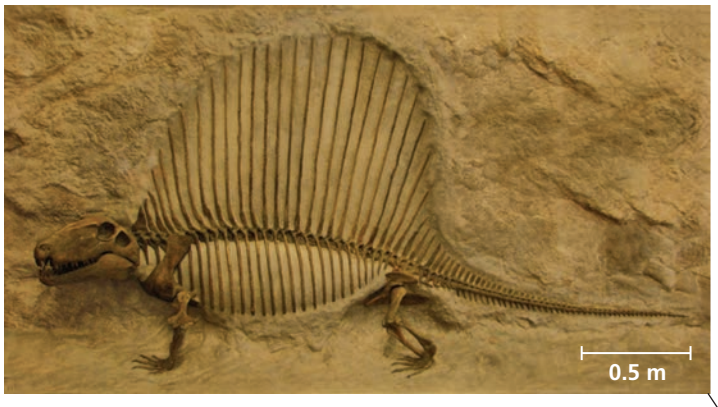
1,500

3,500

▼ *Rhomaleosaurus victor*, a plesiosaur. These large marine reptiles were important predators from 200 million to 66 million years ago.



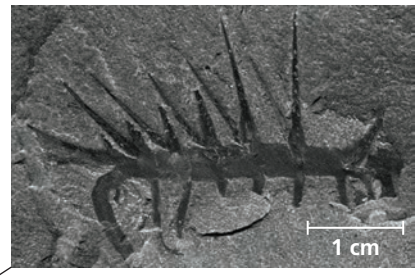
▼ *Dimetrodon*, the largest known carnivore of its day, was more closely related to mammals than to reptiles. The spectacular "sail" on its back may have functioned in temperature regulation or as an ornament that served to attract mates.



▼ *Tiktaalik*, an extinct aquatic organism that is the closest known relative of the four-legged vertebrates that went on to colonize land

▲ *Coccosteus cuspidatus*, a placoderm (fishlike vertebrate) that had a bony shield covering its head and front end

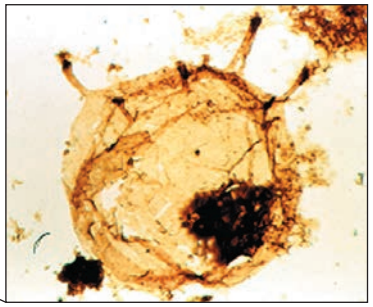
▶ *Hallucigenia*, a member of a morphologically diverse group of animals found in the Burgess Shale fossil bed in the Canadian Rockies



▲ Some prokaryotes bind thin films of sediments together, producing layered rocks called stromatolites. Present-day stromatolites are found in a few shallow marine bays, such as Shark Bay, Australia, shown here.

▲ A section through a fossilized stromatolite

◀ *Dickinsonia costata*, a member of the Ediacaran biota, an extinct group of soft-bodied organisms



▶ *Tappania*, a unicellular eukaryote thought to be either an alga or a fungus

facilitated their fossilization. Even with its limitations, however, the fossil record is a remarkably detailed account of biological change over the vast scale of geologic time. Furthermore, as shown by the recently unearthed fossils of whale ancestors with hind limbs (see Figures 21.19, 21.20, and 25.1), gaps in the fossil record continue to be filled by new discoveries.

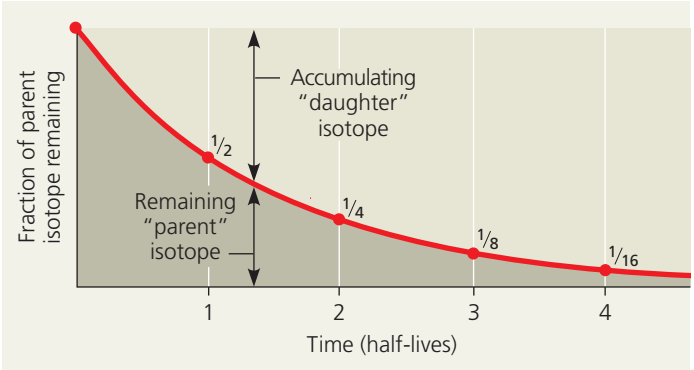
How Rocks and Fossils Are Dated

Fossils are valuable data for reconstructing the history of life, but only if we can determine where they fit in that unfolding story. While the order of fossils in rock strata tells us the sequence in which the fossils were laid down—their relative ages—it does not tell us their actual ages. Examining the relative positions of fossils is like peeling off layers of wallpaper in an old house. You can infer the sequence in which the layers were applied, but not the year each layer was added.

How can we determine the age of a fossil? One of the most common techniques is **radiometric dating**, which is based on the decay of radioactive isotopes (see Concept 2.2). In this process, a radioactive "parent" isotope decays to a "daughter" isotope at a characteristic rate. The rate of decay is expressed by the **half-life**, the time required for 50% of the parent isotope to decay **(Figure 25.6)**. Each type of radioactive isotope has a characteristic half-life, which is not affected by temperature, pressure, or other environmental variables. For example, carbon-14 decays relatively quickly; its half-life is 5,730 years. Uranium-238 decays slowly; its half-life is 4.5 billion years.

Fossils contain isotopes of elements that accumulated in the organisms when they were alive. For example, a living organism contains the most common carbon isotope, carbon-12, as well as a radioactive isotope, carbon-14. When the organism dies, it stops accumulating carbon, and the amount of carbon-12 in its tissues does not change over time. However, the carbon-14 that it contains at the time of death slowly decays into another element, nitrogen-14. Thus, by measuring the ratio of carbon-14 to carbon-12 in a fossil, we can determine the fossil's age. This method works for fossils up to about 75,000 years old; fossils older than that contain too little carbon-14 to be detected with current techniques. Radioactive isotopes with longer half-lives are used to date older fossils.

Figure 25.6 Radiometric dating. In this diagram, each unit of time represents one half-life of a radioactive isotope.



DRAW IT ➤ *Relabel the x-axis of this graph in years to illustrate the radioactive decay of uranium-238 (half-life = 4.5 billion years).*

Determining the age of these older fossils in sedimentary rocks can be challenging. Organisms do not use radioisotopes with long half-lives, such as uranium-238, to build their bones or shells. In addition, sedimentary rocks are often composed of sediments of differing ages. Although we cannot date these older fossils directly, an indirect method can be used to infer the age of fossils that are sandwiched between two layers of volcanic rock. As lava cools into volcanic rock, radioisotopes from the surrounding environment become trapped in the newly formed rock. Some of the trapped radioisotopes have long half-lives, allowing geologists to estimate the ages of ancient volcanic rocks. If two volcanic layers surrounding fossils are found to be 525 million and 535 million years old, for example, then the fossils are roughly 530 million years old.

The Origin of New Groups of Organisms

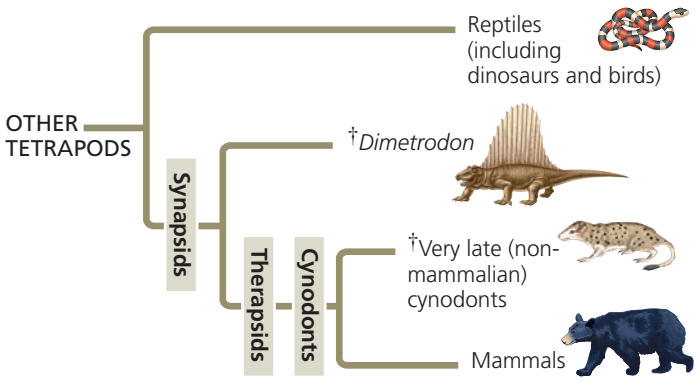
Some fossils provide a detailed look at the origin of new groups of organisms. Such fossils are central to our understanding of evolution; they illustrate how new features arise and how long it takes for such changes to occur. We'll examine one such case here: the origin of mammals.

Along with amphibians and reptiles, mammals belong to the group of animals called *tetrapods* (from the Greek *tetra*, four, and *pod*, foot), named for having four limbs. Mammals have a number of unique anatomical features that fossilize readily, allowing scientists to trace their origin. For example, the lower jaw is composed of one bone (the dentary) in mammals but several bones in other tetrapods. In addition, the lower and upper jaws in mammals hinge between a different set of bones than in other tetrapods. Mammals also have a unique set of three bones that transmit sound in the middle ear, the hammer, anvil, and stirrup, whereas other tetrapods have only one such bone, the stirrup (see Concept 34.6). Finally, the teeth of mammals are differentiated into incisors (for tearing), canines (for piercing), and the multi-pointed premolars and molars (for crushing and grinding). In contrast, the teeth of other tetrapods usually consist of a row of undifferentiated, single-pointed teeth.

As detailed in **Figure 25.7**, the fossil record shows that the unique features of mammalian jaws and teeth evolved gradually over time, in a series of steps. As you study Figure 25.7, bear in mind that it includes just a few examples of the fossil skulls that document the origin of mammals. If all the known fossils in the sequence were arranged by shape and placed side by side, their features would blend smoothly from one group to the next. Some of these fossils would reflect how the features of a group that dominates life today, the mammals, gradually arose in a previously existing group, the cynodonts. Others would reveal side branches on the tree of life—groups of organisms that thrived for millions of years but ultimately left no descendants that survive today.

Figure 25.7 Exploring The Origin of Mammals

Over the course of 120 million years, mammals originated gradually from a group of tetrapods called synapsids. Shown here are a few of the many fossil organisms whose morphological features represent intermediate steps between living mammals and their early synapsid ancestors. The evolutionary context of the origin of mammals is shown in the tree diagram at right (the dagger symbol † indicates extinct lineages).

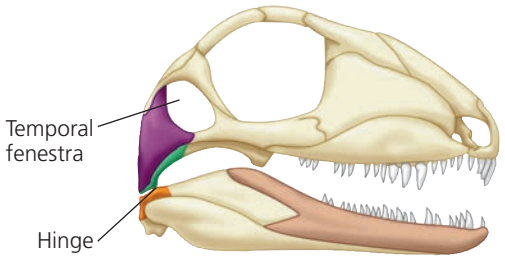


Key to skull bones

Articular
Quadrate
Dentary
Squamosal

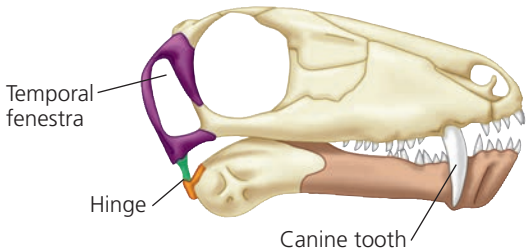
Synapsid (300 mya)

Early synapsids had multiple bones in the lower jaw and single-pointed teeth. The jaw hinge was formed by the articular and quadrate bones. Early synapsids also had an opening called the *temporal fenestra* behind the eye socket. Powerful cheek muscles for closing the jaws probably passed through the temporal fenestra. Over time, this opening enlarged and moved in front of the hinge between the lower and upper jaws, thereby increasing the power and precision with which the jaws could be closed (much as moving a doorknob away from the hinge makes a door easier to close).



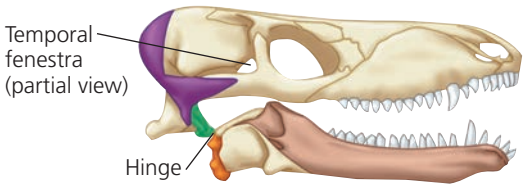
Therapsid (280 mya)

Later, a group of synapsids called therapsids appeared. Therapsids had large dentary bones, long faces, and the first examples of specialized teeth, large canines. These trends continued in a group of therapsids called cynodonts.



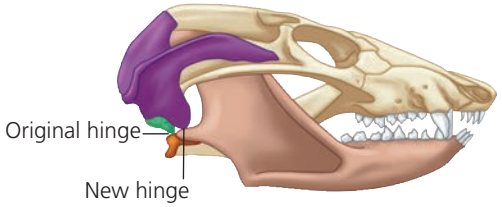
Early cynodont (260 mya)

In early cynodont therapsids, the dentary was the largest bone in the lower jaw, the temporal fenestra was large and positioned forward of the jaw hinge, and teeth with several cusps first appeared (not visible in the diagram). As in earlier synapsids, the jaw had an articular-quadrate hinge.



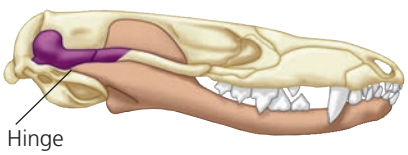
Later cynodont (220 mya)

Later cynodonts had teeth with complex cusp patterns, and their lower and upper jaws hinged in two locations: They retained the original articular-quadrate hinge and formed a new, second hinge between the dentary and squamosal bones. (The temporal fenestra is not visible in this or the below cynodont skull at the angles shown.)



Very late cynodont (195 mya)

In some very late (nonmammalian) cynodonts and early mammals, the original articular-quadrate hinge was lost, leaving the dentary-squamosal hinge as the only hinge between the lower and upper jaws, as in living mammals. The articular and quadrate bones migrated into the ear region (not shown), where they functioned in transmitting sound. In the mammal lineage, these two bones later evolved into the familiar hammer (malleus) and anvil (incus) bones of the ear.



CONCEPT CHECK 25.2

1. Describe an example from the fossil record that shows how life has changed over time.
2. **WHAT IF?** ➤ Your measurements indicate that a fossilized skull you unearthed has a carbon-14/carbon-12 ratio about $1/16$ that of the skulls of present-day animals. What is the approximate age of the fossilized skull?

For suggested answers, see Appendix A.

CONCEPT 25.3

Key events in life's history include the origins of unicellular and multicellular organisms and the colonization of land

The study of fossils has helped geologists establish a **geologic record**: a standard time scale that divides Earth's history into four eons and further subdivisions **(Table 25.1)**. The first three eons—the Hadean, Archaean, and Proterozoic—together lasted about 4 billion years. The Phanerozoic eon, roughly the last half billion years, encompasses most of the time that animals have existed on Earth. It is divided into three eras: the Paleozoic, Mesozoic, and Cenozoic. Each era represents a distinct age in the history of Earth and its life. For example, the Mesozoic era is sometimes called the "age of reptiles" because of its abundance of reptilian fossils, including those of dinosaurs. The boundaries between the eras correspond to major extinction events, when many forms of life disappeared and were replaced by forms that evolved from the survivors.

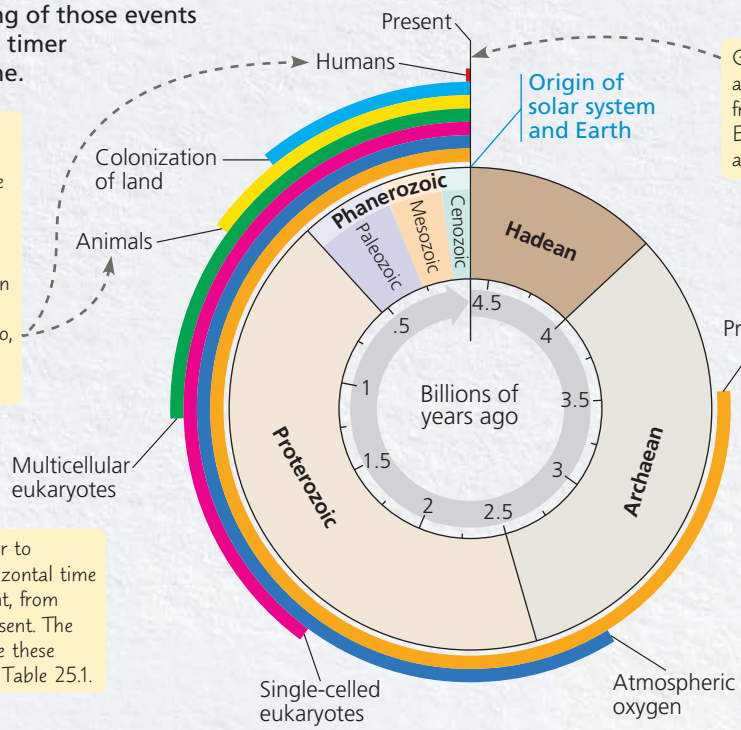
As we've seen, the fossil record provides a sweeping overview of the history of life over geologic time. Here we will focus on a few major events in that history, returning to study the details in Unit Five. **Figure 25.8** will help you visualize how long ago these key events occurred against the vast backdrop of geologic time.

Figure 25.8 Visualizing the Scale of Geologic Time

Geologic time is so vast that it can be difficult to visualize when key events in the history of life on Earth occurred. This figure introduces two common representations that help place the timing of those events in context: a countdown timer and a horizontal time line.

Instructors: Additional questions related to this Visualizing Figure can be assigned in MasteringBiology.

Using the analogy of a timer that begins with the origin of Earth and counts down for one hour, we can relate the relative timing and duration of events that occurred billions of years ago to a familiar time scale. On a one-hour time scale, animals originated about 9 minutes ago, while humans appeared less than 0.2 seconds ago.

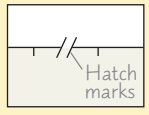


Geologic time is represented here as a timer that, moving clockwise from the top, "counts down" from Earth's origin (4.6 billion years ago) to the present.

1 *Using the analogy of a one-hour countdown timer, when did prokaryotes originate? When did the colonization of land occur?*

This diagram "uncoils" the timer to represent life's history on a horizontal time line. Time runs from left to right, from 4.6 billion years ago to the present. The color-coding will help you relate these diagrams to each other and to Table 25.1.

This diagram shows all of geologic time to scale on an unbroken time line, but it's often necessary to "break" the time line to limit the size of a figure. Hatch marks are often used to represent this interruption; see Figure 25.11 for an example.



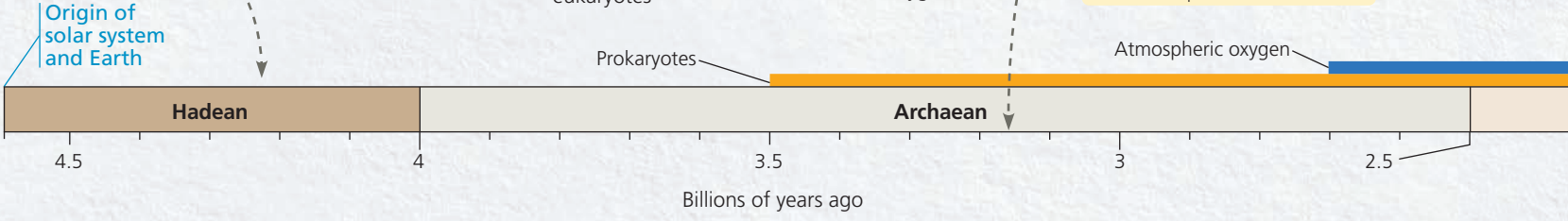
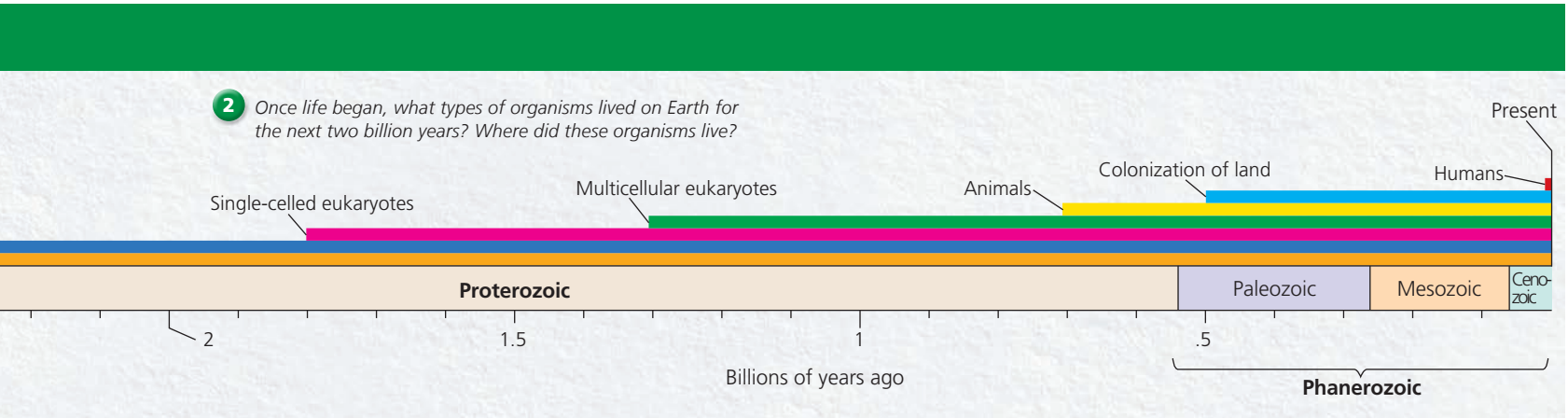
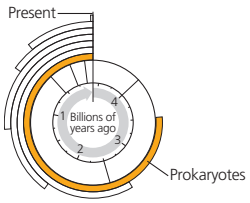


Table 25.1 The Geologic Record

Eons (duration not to scale)	Era	Period	Epoch	Age (Millions of Years Ago)	Some Important Events in the History of Life
Phanerozoic	Cenozoic	Quaternary	Holocene	0.01	Historical time
			Pleistocene	2.6	Ice ages; origin of genus *Homo*
		Neogene	Pliocene	5.3	Appearance of bipedal human ancestors
			Miocene	23	Continued radiation of mammals and angiosperms; earliest direct human ancestors
		Paleogene	Oligocene	34	Origins of many primate groups
			Eocene	56	Angiosperm dominance increases; continued radiation of most present-day mammalian orders
			Paleocene	66	Major radiation of mammals, birds, and pollinating insects
	Mesozoic	Cretaceous		145	Flowering plants (angiosperms) appear and diversify; many groups of organisms, including most dinosaurs, become extinct at end of period
		Jurassic		201	Gymnosperms continue as dominant plants; dinosaurs abundant and diverse
		Triassic		252	Cone-bearing plants (gymnosperms) dominate landscape; dinosaurs evolve and radiate; origin of mammals
	Paleozoic	Permian		299	Radiation of reptiles; origin of most present-day groups of insects; extinction of many marine and terrestrial organisms at end of period
		Carboniferous		359	Extensive forests of vascular plants form; first seed plants appear; origin of reptiles; amphibians dominant
		Devonian		419	Diversification of bony fishes; first tetrapods and insects appear
		Silurian		444	Diversification of early vascular plants
		Ordovician		485	Marine algae abundant; colonization of land by diverse fungi, plants, and animals
		Cambrian		541	Sudden increase in diversity of many animal phyla (Cambrian explosion)
Proterozoic	Neoproterozoic	Ediacaran		635	Diverse algae and soft-bodied invertebrate animals appear
				1,000	
				1,800	Oldest fossils of eukaryotic cells appear
				2,500	
Archaean				2,700	Concentration of atmospheric oxygen begins to increase
				3,500	Oldest fossils of cells (prokaryotes) appear
				4,000	Oldest known rocks on Earth's surface
Hadean				Approx. 4,600	Origin of Earth

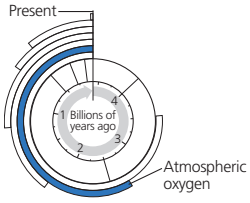


The First Single-Celled Organisms



The earliest direct evidence of life, dating from 3.5 billion years ago, comes from fossilized stromatolites (see Figure 25.5). **Stromatolites** are layered rocks that form when certain prokaryotes bind thin films of sediment together. Stromatolites and other early prokaryotes were Earth's sole inhabitants for about 1.5 billion years. As we will see, these prokaryotes transformed life on our planet.

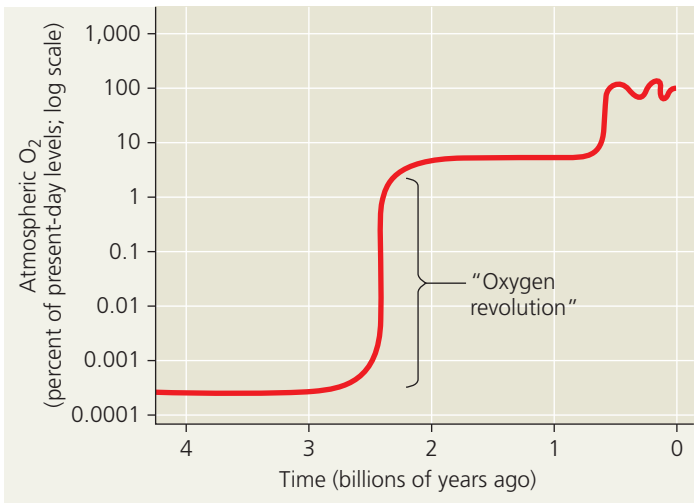
Photosynthesis and the Oxygen Revolution



Most atmospheric oxygen gas (O_2) is of biological origin, produced during the water-splitting step of photosynthesis. When oxygenic photosynthesis first evolved—in photosynthetic prokaryotes— the free O_2 it produced probably dissolved in the surrounding water until it reached a high enough concentration to react with elements dissolved in water, including iron. This would have caused the iron to precipitate as iron oxide, which accumulated as sediments. These sediments were compressed into banded iron formations, red layers of rock containing iron oxide that are a source of iron ore today. Once all of the dissolved iron had precipitated, additional O_2 dissolved in the water until the seas and lakes became saturated with O_2. After this occurred, the O_2 finally began to "gas out" of the water and enter the atmosphere. This change left its mark in the rusting of iron-rich terrestrial rocks, a process that began about 2.7 billion years ago. This chronology implies that bacteria similar to today's cyanobacteria (oxygen-releasing, photosynthetic bacteria) originated before 2.7 billion years ago.

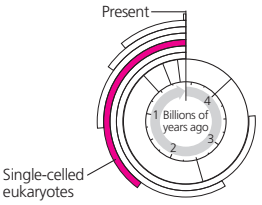
As shown in **Figure 25.9**, the amount of atmospheric O_2 increased gradually from about 2.7 to 2.4 billion years ago, but then shot up relatively rapidly to between 1% and 10% of its present level. This "oxygen revolution" had an enormous impact on life. In some of its chemical forms, oxygen attacks chemical bonds and can inhibit enzymes and damage cells. As a result, the rising concentration of atmospheric O_2 probably doomed many prokaryotic groups. Some species survived in habitats that remained anaerobic, where we find their descendants living today (see Concept 27.4). Among other survivors, diverse adaptations to the changing atmosphere evolved, including cellular respiration, which uses O_2 in the process of harvesting the energy stored in organic molecules.

Figure 25.9 The rise of atmospheric oxygen. Chemical analyses of ancient rocks have enabled this reconstruction of atmospheric oxygen levels during Earth's history.



The rise in atmospheric O_2 levels left a huge imprint on the history of life. A few hundred million years later, another fundamental change occurred: the origin of the eukaryotic cell.

The First Eukaryotes

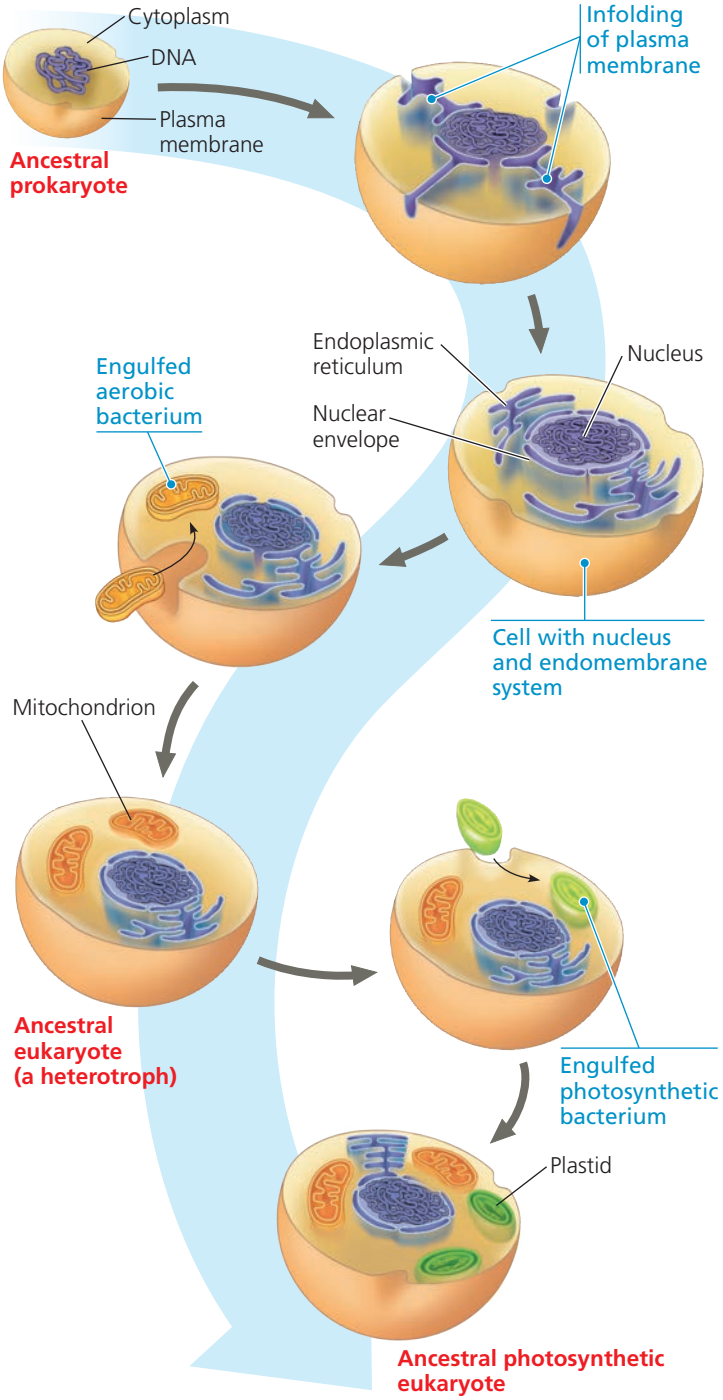


The oldest widely accepted fossils of eukaryotic organisms are 1.8 billion years old. Recall that eukaryotic cells have more complex organization than prokaryotic cells: Eukaryotic cells have a nuclear envelope, mitochondria, endoplasmic reticulum, and other internal structures that prokaryotes lack. Also, unlike prokaryotic cells, eukaryotic cells have a well-developed cytoskeleton, a feature that enables eukaryotic cells to change their shape and thereby surround and engulf other cells.

How did the eukaryotes evolve from their prokaryotic ancestors? Current evidence indicates that the eukaryotes originated by **endosymbiosis** when a prokaryotic cell engulfed a small cell that would evolve into an organelle found in all eukaryotes, the mitochondrion. The small, engulfed cell is an example of an *endosymbiont*, a cell that lives within another cell, called the *host cell*. The prokaryotic ancestor of the mitochondrion probably entered the host cell as undigested prey or an internal parasite. Though such a process may seem unlikely, scientists have directly observed cases in which endosymbionts that began as prey or parasites developed a mutually beneficial relationship with the host in as little as five years.

By whatever means the relationship began, we can hypothesize how the symbiosis could have become beneficial. For example, in a world that was becoming increasingly aerobic, a host that was itself an anaerobe would have benefited from endosymbionts that could make use of the oxygen. Over time, the host and endosymbionts would have become a single organism, its parts inseparable. Although all eukaryotes have mitochondria or remnants of these organelles, they do not all have plastids (a general term for chloroplasts and related organelles). Thus, the **serial endosymbiosis** hypothesis supposes that mitochondria evolved before plastids through a sequence of endosymbiotic events. As shown in **Figure 25.10**, both

Figure 25.10 A hypothesis for the origin of mitochondria and plastids through serial endosymbiosis. The proposed host was an archaean or a close relative of the archaeans. The proposed ancestors of mitochondria were aerobic, heterotrophic bacteria, while the proposed ancestors of plastids were photosynthetic bacteria. In this figure, the arrows represent change over evolutionary time.



MB Figure Walkthrough

mitochondria and plastids are thought to have descended from bacterial cells. The original host—the cell that engulfed the bacterium whose descendants gave rise to the mitochondrion—is thought to have been an archaean or a close relative of the archaeans.

A great deal of evidence supports the endosymbiotic origin of mitochondria and plastids:

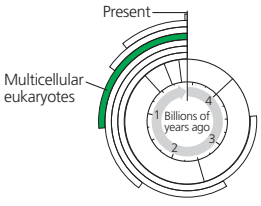
- The inner membranes of both organelles have enzymes and transport systems that are homologous to those found in the plasma membranes of living bacteria.
- Mitochondria and plastids replicate by a splitting process that is similar to that of certain bacteria. In addition, each of these organelles contains circular DNA molecules that, like the chromosomes of bacteria, are not associated with histones or large amounts of other proteins.
- As might be expected of organelles descended from free-living organisms, mitochondria and plastids also have the cellular machinery (including ribosomes) needed to transcribe and translate their DNA into proteins.
- Finally, in terms of size, RNA sequences, and sensitivity to certain antibiotics, the ribosomes of mitochondria and plastids are more similar to bacterial ribosomes than they are to the cytoplasmic ribosomes of eukaryotic cells.

In Chapter 28, we'll return to the origin of eukaryotes, focusing on what genomic data have revealed about the prokaryotic lineages that gave rise to the host and endosymbiont cells.

The Origin of Multicellularity

An orchestra can play a greater variety of musical compositions than a violin soloist can; the increased complexity of the orchestra makes more variations possible. Likewise, the origin of structurally complex eukaryotic cells sparked the evolution of greater morphological diversity than was possible for the simpler prokaryotic cells. After the first eukaryotes appeared, a great range of unicellular forms evolved, giving rise to the diversity of single-celled eukaryotes that continue to flourish today. Another wave of diversification also occurred: Some single-celled eukaryotes gave rise to multicellular forms, whose descendants include a variety of algae, plants, fungi, and animals.

Early Multicellular Eukaryotes

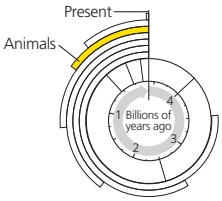


The oldest known fossils of multicellular eukaryotes that can be resolved taxonomically are of relatively small red algae that lived 1.2 billion years ago; even older fossils, dating to 1.8 billion years ago, may also be of small, multicellular eukaryotes. Larger and more diverse multicellular eukaryotes do not appear in the fossil record until about 600 million years ago (see Figure 25.5). These fossils, referred to as the Ediacaran biota, were of soft-bodied organisms—some over 1 m long—that lived from 635 to 541 million years ago. The Ediacaran biota included both algae and animals, along with various organisms of unknown taxonomic affinity.

The rise of large eukaryotes in the Ediacaran period represents an enormous change in the history of life. Before

that time, Earth was a microbial world: Its only inhabitants were single-celled prokaryotes and eukaryotes, along with an assortment of microscopic, multicellular eukaryotes. As the diversification of the Ediacaran biota came to a close about 541 million years ago, the stage was set for another, even more spectacular burst of evolutionary change: the "Cambrian explosion."

The Cambrian Explosion

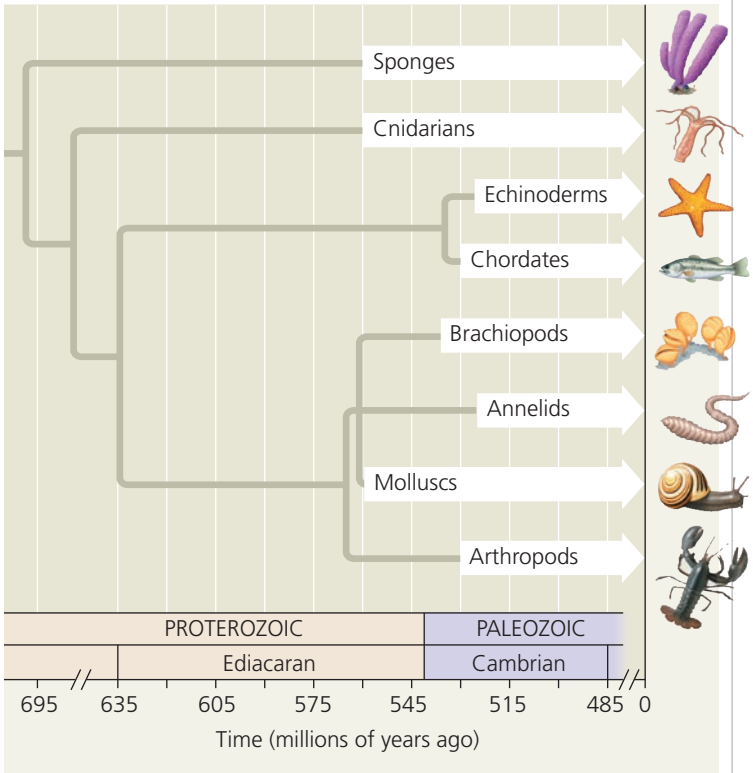


Many present-day animal phyla appear suddenly in fossils formed 535–525 million years ago, early in the Cambrian period. This phenomenon is referred to as the **Cambrian explosion**. Fossils of several animal groups—sponges, cnidarians (sea anemones and their relatives), and molluscs (snails, clams, and their relatives)—appear in even older rocks dating from the late Proterozoic **(Figure 25.11)**.

Prior to the Cambrian explosion, all large animals were soft-bodied. The fossils of large pre-Cambrian animals reveal little evidence of predation. Instead, these animals appear to have been grazers (feeding on algae), filter feeders, or scavengers, not hunters. The Cambrian explosion changed all of that. In a relatively short period of time (10 million years), predators over 1 m in length emerged that had claws and other features for capturing prey; simultaneously, new defensive adaptations, such as sharp spines and heavy body armor, appeared in their prey (see Figure 25.5).

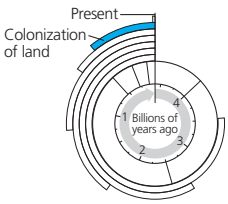
Although the Cambrian explosion had an enormous impact on life on Earth, it appears that many animal phyla originated long before that time. Recent DNA analyses suggest that sponges had evolved by 700 million years ago; such analyses also indicate that the common ancestor of arthropods, chordates, and other animal phyla that radiated during the Cambrian explosion lived 670 million years ago. Researchers have unearthed 710-million-year-old sediments containing steroids indicative of a particular group of sponges—a finding that supports the molecular data. In contrast, the oldest fossil assigned to an extant animal phylum is that of the mollusc *Kimberella*, which lived 560 million years ago. Overall, molecular and fossil data indicate that the Cambrian explosion had a "long fuse"—at least 25 million years long based on the age of *Kimberella* fossils, and over 100 million years long based on some DNA analyses.

Figure 25.11 Appearance of selected animal groups. The white bars indicate earliest appearances of these animal groups in the fossil record.



VISUAL SKILLS ➤ *Circle the branch point that represents the most recent common ancestor of chordates and annelids. What is a minimum estimate of that ancestor's age?*

The Colonization of Land

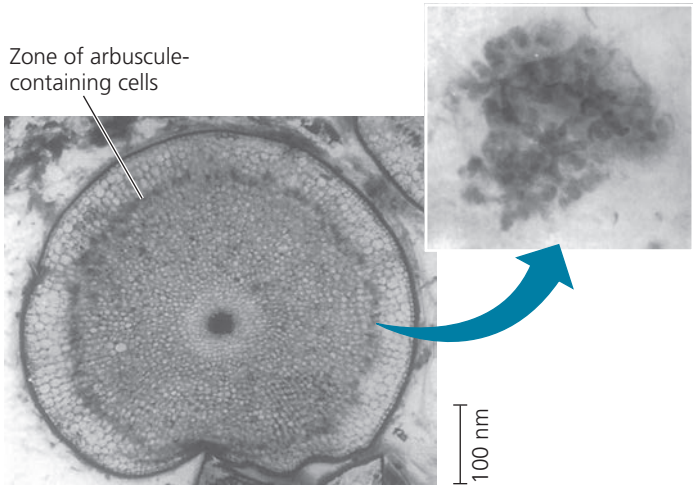


The colonization of land was another milestone in the history of life. There is fossil evidence that cyanobacteria and other photosynthetic prokaryotes coated damp terrestrial surfaces well over a billion years ago. However, larger forms of life, such as fungi, plants, and animals, did not begin to colonize land until about 500 million years ago. This gradual evolutionary venture out of aquatic environments was associated with adaptations that made it possible to reproduce on land and that helped prevent dehydration. For example, many land plants today have a vascular system for transporting materials internally and a waterproof coating of wax on their leaves that slows the loss of water to the air. Early signs of these adaptations were present 420 million years ago, at which time small plants (about 10 cm high) existed that had a vascular system but lacked true roots or leaves. By 40 million years later, plants had diversified greatly and included reeds and treelike plants with true roots and leaves.

Plants appear to have colonized land in the company of fungi. Even today, the roots of most plants are associated with fungi that aid in the absorption of water and minerals from the soil (see Concept 31.1). These root fungi (or *mycorrhizae*), in turn, obtain their organic nutrients from the plants. Such mutually beneficial associations of plants and fungi are evident in some of the oldest fossilized plants, dating this relationship back to the early spread of life onto land **(Figure 25.12)**.

Although many animal groups are now represented in terrestrial environments, the most widespread and diverse

Figure 25.12 An ancient symbiosis. This 405-million-year-old fossil stem (cross section) documents mycorrhizae in the early land plant *Aglaophyton major*. The inset shows an enlarged view of a cell containing a branched fungal structure called an arbuscule; the fossil arbuscule resembles those seen in plant cells today.



land animals are arthropods (particularly insects and spiders) and tetrapods. Arthropods were among the first animals to colonize land, roughly 450 million years ago. The earliest tetrapods found in the fossil record lived about 365 million years ago and appear to have evolved from a group of lobe-finned fishes (see Concept 34.3). Tetrapods include humans, although we are late arrivals on the scene. The human lineage diverged from other primates around 6–7 million years ago, and our species originated only about 195,000 years ago. If the clock of Earth's history were rescaled to represent an hour, humans appeared less than 0.2 second ago.

Interview with Geerat Vermeij: What fossils reveal about the history of life

CONCEPT CHECK 25.3

1. The first appearance of free oxygen in the atmosphere likely triggered a massive wave of extinctions among the prokaryotes of the time. Why?
2. What evidence supports the hypothesis that mitochondria preceded plastids in the evolution of eukaryotic cells?
3. **WHAT IF?** ➤ What would a fossil record of life today look like?

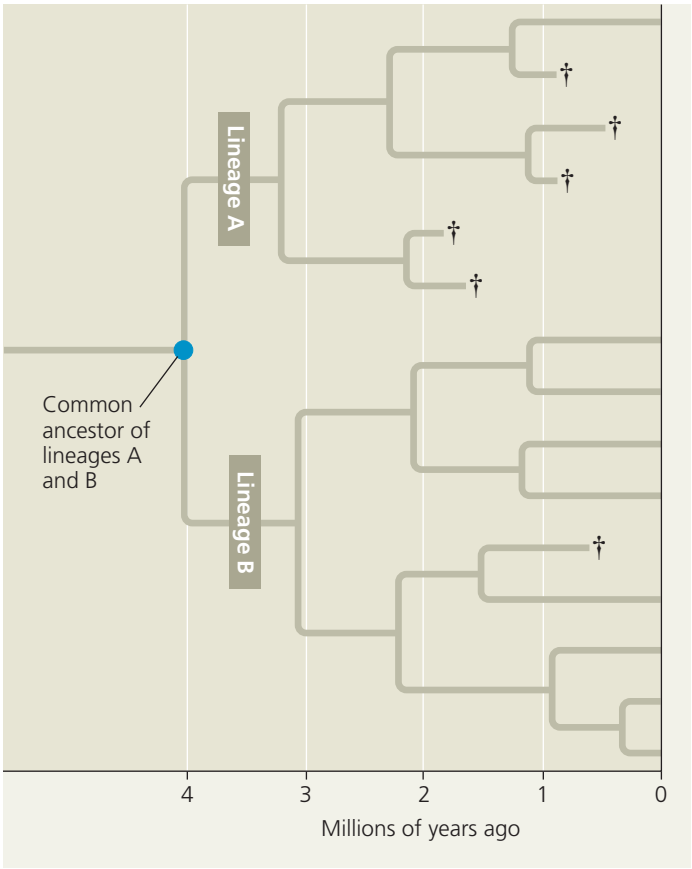
For suggested answers, see Appendix A.

CONCEPT 25.4

The rise and fall of groups of organisms reflect differences in speciation and extinction rates

From its beginnings, life on Earth has been marked by the rise and fall of groups of organisms. Anaerobic prokaryotes originated, flourished, and then declined as the oxygen content of the atmosphere rose. Billions of years later, the first tetrapods emerged from the sea, giving rise to several major new groups of organisms. One of these, the amphibians, went on to dominate life on land for 100 million years, until other tetrapods (including dinosaurs and, later, mammals) replaced them as the dominant terrestrial vertebrates.

Figure 25.13 How speciation and extinction affect diversity. The species diversity of an evolutionary lineage will increase when more new member species originate than are lost to extinction. In this hypothetical example, by 2 million years ago both lineage A and lineage B have given rise to four species, and no species have become extinct. Over the next 2 million years, however, lineage A experiences higher extinction rates than does lineage B (extinct species are denoted by a dagger symbol, †). As a result, after 4 million years (that is, by time 0), lineage A contains only one species, while lineage B contains eight species.



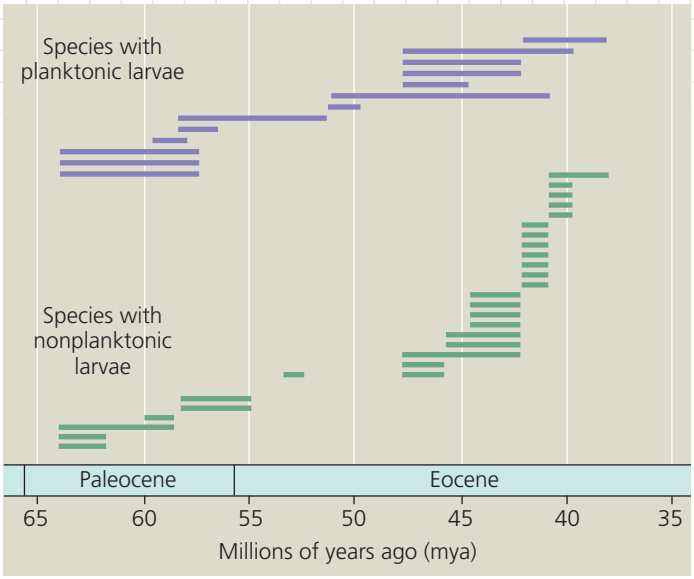
The rise and fall of these and other major groups of organisms have shaped the history of life. Narrowing our focus, we can also see that the rise or fall of any particular group is related to the speciation and extinction rates of its member species **(Figure 25.13)**. Just as a population increases in size when there are more births than deaths, the rise of a group of organisms occurs when more new species are produced than are lost to extinction. The reverse occurs when a group is in decline. In the **Scientific Skills Exercise**, you will interpret data from the fossil record about changes in a group of snail species in the early Paleogene period. Such changes in the fates of groups of organisms have been influenced by large-scale processes such as plate tectonics, mass extinctions, and adaptive radiations.

SCIENTIFIC SKILLS EXERCISE

Estimating Quantitative Data from a Graph and Developing Hypotheses

Do Ecological Factors Affect Evolutionary Rates? Researchers studied the fossil record to investigate whether differing modes of larval dispersal might explain species longevity within one taxon of marine snails, the family Volutidae. Some of the snail species had nonplanktonic larvae: They developed directly into adults without a swimming stage. Other species had planktonic larvae: They had a swimming stage and could disperse very long distances. The adults of these planktonic species tended to have broad geographic distributions, whereas nonplanktonic species tended to be more isolated.

How the Research Was Done The researchers studied the stratigraphic distribution of volutes in outcrops of sedimentary rocks located along North America's Gulf coast. These rocks, which formed from 66 to 37 million years ago, early in the Paleogene period, are an excellent source of well-preserved snail fossils. The researchers were able to classify each fossil species of volute snail as having planktonic or nonplanktonic larvae based on features of the earliest formed whorls of the snail's shell. Each bar in the graph shows how long one species of snail persisted in the fossil record.



INTERPRET THE DATA

1. You can estimate quantitative data (fairly precisely) from a graph. The first step is to obtain a conversion factor by measuring along an axis that has a scale. In this case, 25 million years (my; from 60 to 35 million years ago [mya] on the *x*-axis) is represented by a distance of 7.0 cm. This yields a conversion factor (a ratio) of 25 my/7.0 cm = 3.6 my/cm. To estimate the time period represented by a horizontal bar on this graph, measure the length of that bar in centimeters and multiply that measurement by the conversion factor, 3.6 my/cm. For example, a bar that measures 1.1 cm on the graph represents a persistence time of 1.1 cm × 3.6 my/cm = 4 million years.
2. Calculate the mean (average) persistence times for species with planktonic larvae and species with nonplanktonic larvae.
3. Count the number of new species that form in each group beginning at 60 mya (the first three species in each group were present around 64 mya, the first time period sampled, so we don't know when those species first appear in the fossil record).
4. Propose a hypothesis to explain the differences in longevity of snail species with planktonic and nonplanktonic larvae.

Data from T. A. Hansen, Larval dispersal and species longevity in Lower Tertiary gastropods, *Science* 199:885–887 (1978). Reprinted with permission from AAAS.

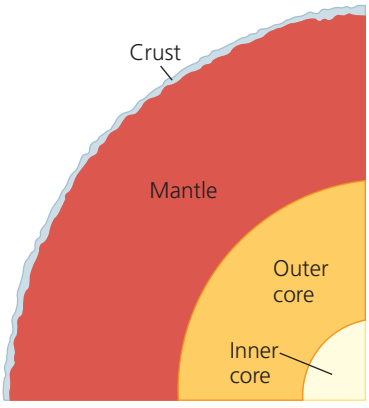
Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Plate Tectonics

If photographs of Earth were taken from space every 10,000 years and spliced together to make a movie, it would show something many of us find hard to imagine: The seemingly "rock solid" continents we live on move over time. Over the past billion years, there have been three occasions (1 billion, 600 million, and 250 million years ago) when most of the landmasses of Earth came together to form a supercontinent, then later broke apart. Each time, this breakup yielded a different configuration of continents. Based on the directions in which the continents are moving today, some geologists have estimated that a new supercontinent will form roughly 250 million years from now.

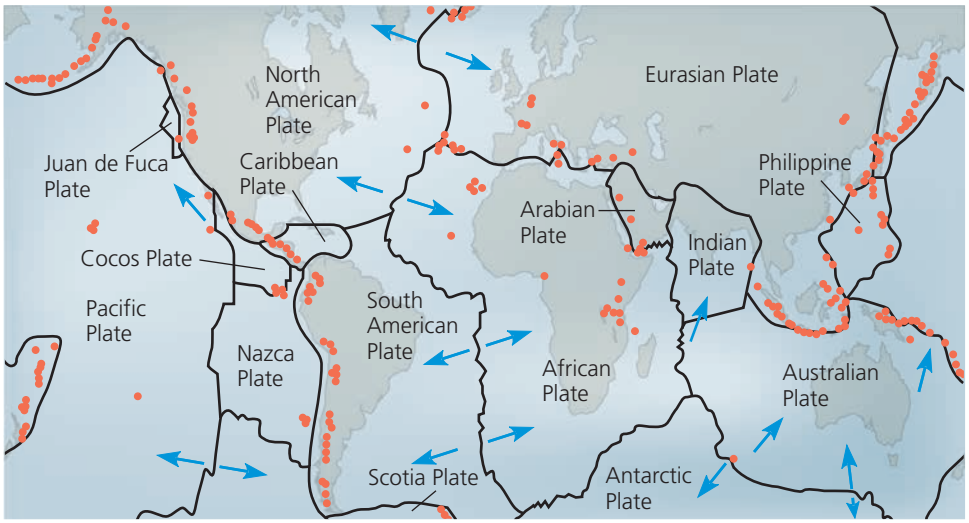
According to the theory of **plate tectonics**, the continents are part of great plates of Earth's crust that essentially float on the hot, underlying portion of the mantle **(Figure 25.14)**. Movements in the mantle cause the plates to move over time in a process called *continental drift*. Geologists can measure the rate at which the plates are moving now, usually only a few centimeters per year. They can also infer the past locations of the continents using the magnetic signal recorded in rocks at the time of their formation. This method works because as a continent shifts its position over time, the direction of magnetic north recorded in its newly formed rocks also changes.

Figure 25.14 Cutaway view of Earth. The thickness of the crust is exaggerated here.



Earth's major tectonic plates are shown in **Figure 25.15**. Many important geologic processes, including the formation of mountains and islands, occur at plate boundaries. In some cases, two plates are moving away from each other, as are the North American and Eurasian

Figure 25.15 Earth's major tectonic plates. The arrows indicate direction of movement. The reddish orange dots represent zones of violent tectonic activity.



HHMI Animation: Plate Tectonics

plates, which are currently drifting apart at a rate of about 2 cm per year. In other cases, two plates are sliding past each other, forming regions where earthquakes are common. California's infamous San Andreas Fault is part of a border where two plates slide past each other. In still other cases, two plates collide, producing violent upheavals and forming new mountains along the plate boundaries. One spectacular example of this occurred 45 million years ago, when the Indian plate crashed into the Eurasian plate, starting the formation of the Himalayan mountains.

HHMI Video: Animated Life: Pangea, Wegener, and Continental Drift

Consequences of Continental Drift

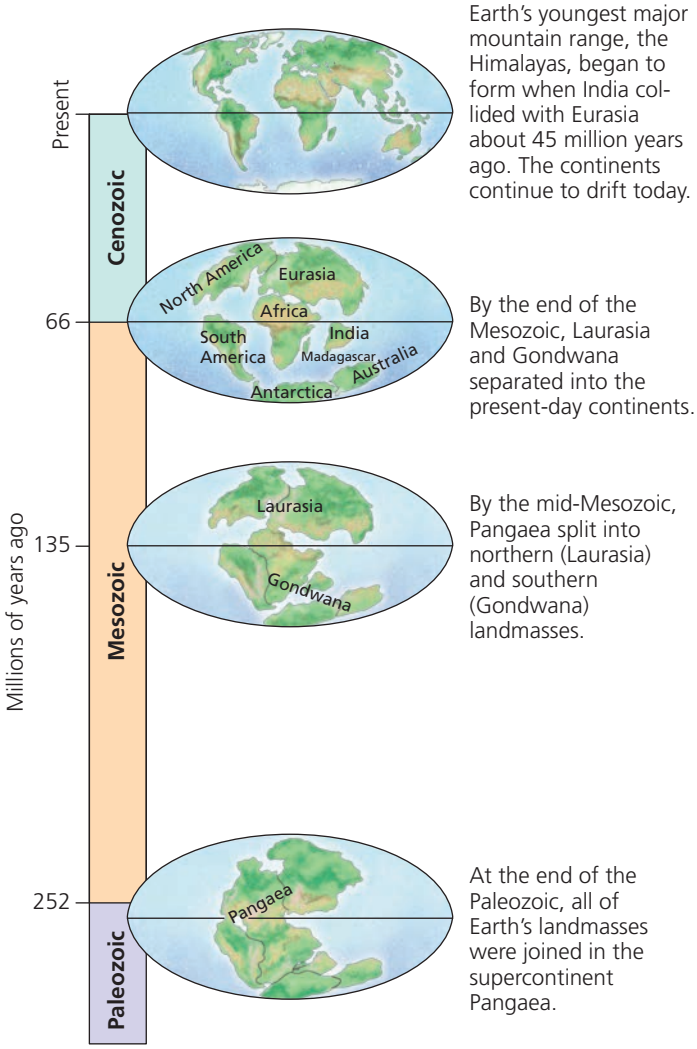
Plate movements rearrange geography slowly, but their cumulative effects are dramatic. In addition to reshaping the physical features of our planet, continental drift also has a major impact on life on Earth.

One reason for this is that continental drift alters the habitats in which organisms live. Consider the changes shown in **Figure 25.16**. About 250 million years ago, plate movements brought previously separated landmasses together into a supercontinent named **Pangaea**. Ocean basins became deeper, which drained shallow coastal seas. At that time, as now, most marine species inhabited shallow waters, and the formation of Pangaea destroyed much of that habitat. Pangaea's interior was cold and dry, probably an even more severe environment than that of central Asia today. Overall, the formation of Pangaea greatly altered the physical environment and climate, which drove some species to extinction and provided new opportunities for groups of organisms that survived the crisis.

Organisms are also affected by the climate change that results when a continent shifts its location. The southern tip of Labrador, Canada, for example, once was located in the tropics but has moved 40° to the north over the last 200 million years. When faced with the changes in climate that such shifts in position entail, organisms adapt, move to a new location, or become extinct (this last outcome occurred for many organisms stranded on Antarctica, which separated from Australia 40 million years ago).

Continental drift also promotes allopatric speciation on a grand scale. When supercontinents break apart, regions that once were connected become isolated. As the continents drifted apart over the last 200 million years, each became a separate evolutionary arena, with lineages of plants and animals that diverged from those on other continents.

Figure 25.16 The history of continental drift during the Phanerozoic eon.



VISUAL SKILLS > *Is the Australian plate's current direction of movement (see Figure 25.15) similar to the direction it traveled over the past 66 million years?*

Finally, continental drift can help explain puzzles about the geographic distribution of extinct organisms, such as why fossils of the same species of Permian freshwater reptiles have been discovered in both Brazil and the West African nation of Ghana. These two parts of the world, now separated by 3,000 km of ocean, were joined together when these reptiles were living. Continental drift also explains much about the current distributions of organisms, such as why Australian fauna and flora contrast so sharply with those of the rest of the world. Marsupial mammals fill ecological roles in Australia analogous to those filled by eutherians (placental mammals) on other continents (see Figure 21.18). Fossil evidence suggests that marsupials originated in what is now Asia and reached Australia via South America and Antarctica while the continents were still joined. The subsequent breakup of the southern continents set Australia "afloat," like a giant raft of marsupials. In Australia, marsupials diversified, and the few eutherians that lived there became extinct; on other continents, most marsupials became extinct, and the eutherians diversified.

Mass Extinctions

The fossil record shows that the overwhelming majority of species that ever lived are now extinct. A species may become extinct for many reasons. Its habitat may have been destroyed, or its environment may have changed in a manner unfavorable to the species. For example, if ocean temperatures fall by even a few degrees, species that are otherwise well adapted may perish. Even if physical factors in the environment remain stable, biological factors may change—the origin of one species can spell doom for another.

Although extinction occurs regularly, at certain times disruptive changes to the global environment have caused the rate of extinction to increase dramatically. The result is a **mass extinction**, in which large numbers of species become extinct worldwide.

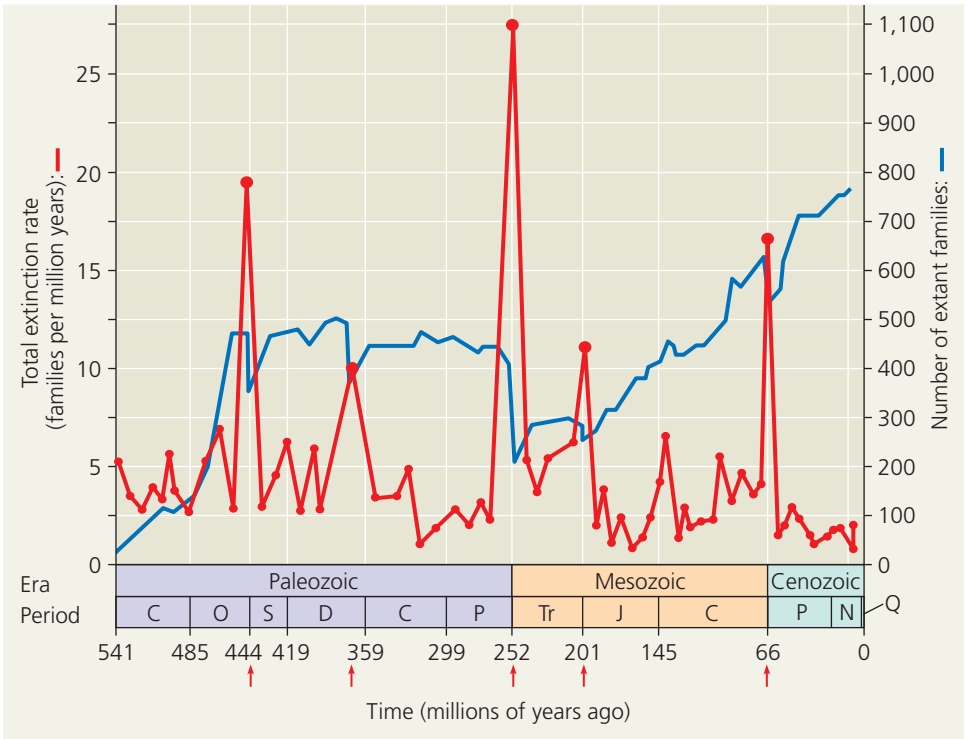
The "Big Five" Mass Extinction Events

Five mass extinctions are documented in the fossil record over the past 500 million years **(Figure 25.17)**. These events are particularly well documented for the decimation of hard-bodied animals that lived in shallow seas, the organisms for which the fossil record is most complete. In each mass extinction, 50% or more of marine species became extinct.

Two mass extinctions—the Permian and the Cretaceous—have received the most attention. The Permian mass extinction, which defines the boundary between the Paleozoic and Mesozoic eras (252 million years ago), claimed about 96% of marine animal species and drastically altered life in the ocean. Terrestrial life was also affected. For example, 8 out of 27 known orders of insects were wiped out. This mass extinction occurred in less than 500,000 years, possibly in just a few thousand years—an instant in the context of geologic time.

The Permian mass extinction occurred during the most extreme episode of volcanism in the past 500 million years. Geologic data indicate that 1.6 million km^2 (roughly half the size of Western Europe) in Siberia was covered with lava hundreds of meters thick. The eruptions are thought to have produced enough carbon dioxide to warm the global climate by an estimated 6°C, harming many temperature-sensitive species. The rise in atmospheric CO_2 levels would also have led to ocean acidification, thereby reducing the availability of calcium carbonate, which is required by reef-building corals and many shell-building species (see Figure 3.12). The explosions would also have added nutrients such as phosphorus to marine ecosystems, stimulating the growth of microorganisms. Upon their deaths, these microorganisms would have provided food for bacterial decomposers. Bacteria use oxygen as they decompose the bodies of dead organisms, thus causing oxygen concentrations to drop. This would have harmed

Figure 25.17 Mass extinction and the diversity of life. The five generally recognized mass extinction events, indicated by red arrows, represent peaks in the extinction rate of marine animal families (red line and left vertical axis). These mass extinctions interrupted the overall increase, over time, in the number of extant families of marine animals (blue line and right vertical axis).



INTERPRET THE DATA *As mentioned in the text, 96% of marine animal species became extinct in the Permian mass extinction. Explain why the blue curve shows only a 50% drop at that time.*

oxygen-breathers and promoted the growth of anaerobic bacteria that emit a poisonous metabolic by-product, hydrogen sulfide (H_2S) gas. Overall, the volcanic eruptions may have triggered a series of catastrophic events that together resulted in the Permian mass extinction.

The Cretaceous mass extinction occurred 66 million years ago. This event extinguished more than half of all marine species and eliminated many families of terrestrial plants and animals, including all dinosaurs (except birds, which are members of the same group; see Figure 34.25). One clue to a possible cause of the Cretaceous mass extinction is a thin layer of clay enriched in iridium that dates to the time of the mass extinction. Iridium is an element that is very rare on Earth but common in many of the meteorites and other extraterrestrial objects that occasionally fall to Earth. As a result, researchers proposed that this clay is fallout from a huge cloud of debris that billowed into the atmosphere when an asteroid or large comet collided with Earth. This cloud would have blocked sunlight and severely disturbed the global climate for several months.

Is there evidence of such an asteroid or comet? Research has focused on the Chicxulub crater, a 66-million-year-old scar beneath sediments off the coast of Mexico **(Figure 25.18)**.

Figure 25.18 A trauma for Cretaceous life. Beneath the Caribbean Sea, the 66-million-year-old Chicxulub crater measures 180 km across. The horseshoe shape of the crater and the pattern of debris in sedimentary rocks indicate that an asteroid or comet struck at a low angle from the southeast. This drawing represents the impact and its immediate effect: a cloud of hot vapor and debris that could have killed many of the plants and animals in North America within hours.

The crater is the right size to have been caused by an object with a diameter of 10 km. Critical evaluation of this and other hypotheses for mass extinctions continues.

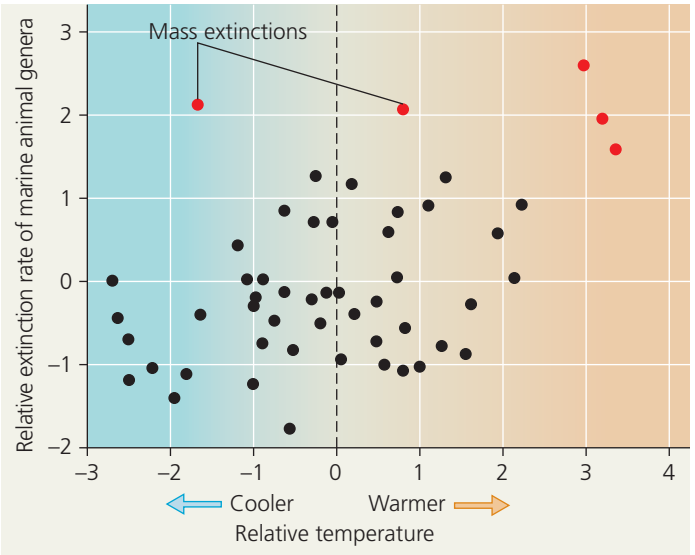
HHMI Video: The Day the Mesozoic Died

Is a Sixth Mass Extinction Under Way?

As you will read further in Concept 56.1, human actions, such as habitat destruction, are modifying the global environment to such an extent that many species are threatened with extinction. More than 1,000 species have become extinct in the last 400 years. Scientists estimate that this rate is 100 to 1,000 times the typical background rate seen in the fossil record. Is a sixth mass extinction now in progress?

This question is difficult to answer, in part because it is hard to document the total number of extinctions occurring today. Tropical rain forests, for example, harbor many undiscovered species. As a result, destroying tropical forest may drive species to extinction before we even learn of their existence. Such uncertainties make it hard to assess the full extent of the current extinction crisis. Even so, it is clear that losses to date have not reached those of the "big five" mass extinctions, in which large percentages of Earth's species became extinct. This does not in any way discount the seriousness of today's situation. Monitoring programs show that many species are declining at an alarming rate due to habitat loss, introduced species, overharvesting, and other factors. Recent studies on a variety of organisms, including lizards, pine trees, and polar bears, suggest that climate change may hasten some of these declines. The fossil record also highlights the potential importance of climate change: Over the last 500 million years, extinction rates have tended to increase when global temperatures were high **(Figure 25.19)**.

Figure 25.19 Fossil extinctions and temperature. Extinction rates increased when global temperatures were high. Temperatures were estimated using ratios of oxygen isotopes and converted to an index in which 0 is the overall average temperature.



Overall, the evidence suggests that unless dramatic actions are taken, a sixth, human-caused mass extinction is likely to occur within the next few centuries or millennia.

Consequences of Mass Extinctions

Mass extinctions have significant and long-term effects. By eliminating large numbers of species, a mass extinction can reduce a thriving and complex ecological community to a pale shadow of its former self. And once an evolutionary lineage disappears, it cannot reappear. The course of evolution is changed forever. Consider what would have happened if the early primates living 66 million years ago had died out in the Cretaceous mass extinction. Humans would not exist, and life on Earth would differ greatly from what it is today.

The fossil record shows that it typically takes 5–10 million years for the diversity of life to recover to previous levels after a mass extinction. In some cases, it has taken much longer than that: It took about 100 million years for the number of marine families to recover after the Permian mass extinction (see Figure 25.17). These data have sobering implications. If current trends continue and a sixth mass extinction occurs, it will take millions of years for life on Earth to recover.

Mass extinctions can also alter ecological communities by changing the types of organisms residing there. For example, after the Permian and Cretaceous mass extinctions, the percentage of marine organisms that were predators increased substantially **(Figure 25.20)**. A rise in the number of predators can increase both the risks faced by prey and the competition among predators for food. In addition, mass extinctions can curtail lineages with novel and advantageous features. For example, in the late Triassic period, a group of gastropods (snails and their relatives) arose that could drill through the shells of bivalves (such as clams) and feed on the animals inside. Although shell drilling provided access to a new and abundant source of food, this newly formed group was wiped out during the mass extinction at the end of the Triassic (about 200 million years ago). Another 120 million years passed before another group of gastropods (the oyster drills) exhibited the ability to drill through shells. As their predecessors might have done if they had not originated at an unfortunate time, oyster drills have since diversified into many new species. Finally, by eliminating so many species, mass extinctions can pave the way for adaptive radiations, in which new groups of organisms proliferate.

Adaptive Radiations

The fossil record shows that the diversity of life has increased over the past 250 million years (see blue line in Figure 25.17). This increase has been fueled by **adaptive radiations**, periods of evolutionary change in which groups of organisms form many new species whose adaptations allow them to fill different ecological roles, or niches, in their communities. Large-scale adaptive radiations occurred after each of the big five mass extinctions, when survivors became adapted to the many vacant ecological niches. Adaptive radiations have also occurred in groups of organisms that possessed major evolutionary innovations, such as seeds or armored body coverings, or that colonized regions in which they faced little competition from other species.

Worldwide Adaptive Radiations

Fossil evidence indicates that mammals underwent a dramatic adaptive radiation after the extinction of terrestrial dinosaurs 66 million years ago **(Figure 25.21)**. Although mammals originated about 180 million years ago, the mammal fossils older than 66 million years are mostly small and not morphologically diverse. Many species appear to have been nocturnal based on their large eye sockets, similar to those in living nocturnal mammals. A few early mammals were intermediate in size, such as *Repenomamus giganticus*, a 1-m-long predator that lived 130 million years ago—but none approached the size of many dinosaurs. Early mammals may have been restricted in size and diversity because they were eaten or outcompeted by the larger and more diverse dinosaurs. With the disappearance of the dinosaurs (except for birds), mammals expanded greatly in both diversity and size, filling the ecological roles once occupied by terrestrial dinosaurs.

The history of life has also been greatly altered by radiations in which groups of organisms increased in diversity as they came to play entirely new ecological roles in their communities. As we'll explore

Figure 25.20 Mass extinctions and ecology. The Permian and Cretaceous mass extinctions (indicated by red arrows) altered the ecology of the oceans by increasing the percentage of marine genera that were predators.

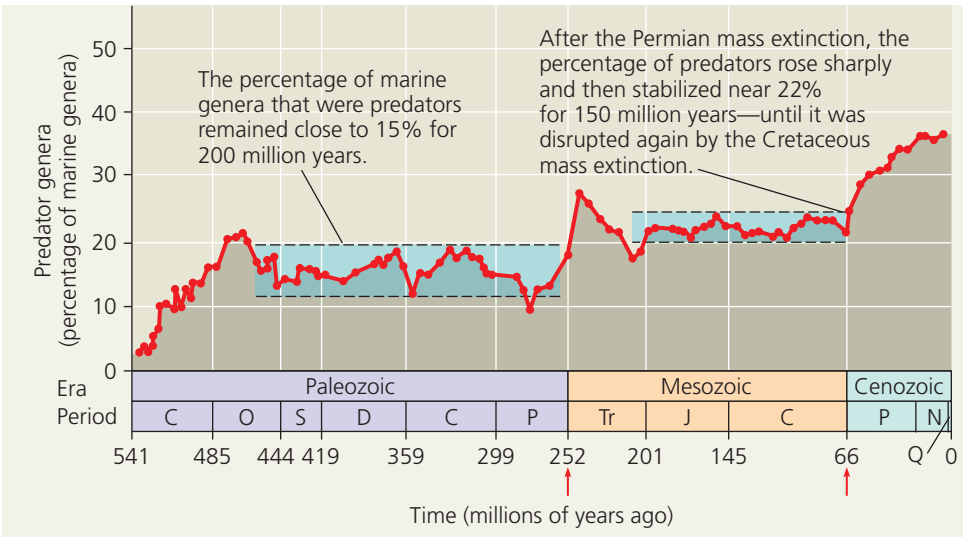
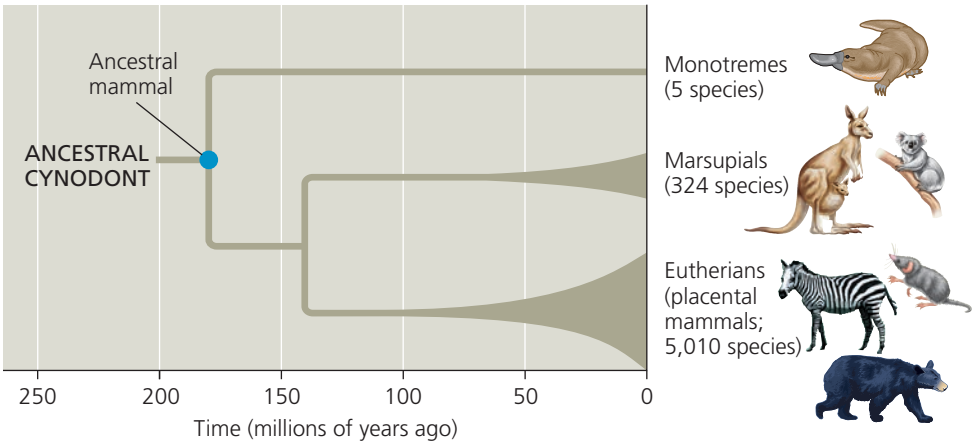


Figure 25.21 Adaptive radiation of mammals.



in later chapters, examples include the rise of photosynthetic prokaryotes, the evolution of large predators in the Cambrian explosion, and the radiations following the colonization of land by plants, insects, and tetrapods. Each of these last three radiations was associated with major evolutionary innovations that facilitated life on land. The radiation of land plants, for example, was associated with key adaptations, such as stems that support plants against gravity and a waxy coat that protects leaves from water loss. Finally, organisms that arise in an adaptive radiation can serve as a new source of food for still other organisms. In fact, the diversification of land plants stimulated a series of adaptive radiations in insects that ate or pollinated plants, one reason that insects are the most diverse group of animals on Earth today.

Regional Adaptive Radiations

Striking adaptive radiations have also occurred over more limited geographic areas. Such radiations can be initiated when a few organisms make their way to a new, often distant location in which they face relatively little competition from other organisms. The Hawaiian archipelago is one of the world's great showcases of this type of adaptive radiation **(Figure 25.22)**. Located about 3,500 km from the nearest continent, the volcanic islands are progressively older as one follows the chain toward the northwest; the youngest

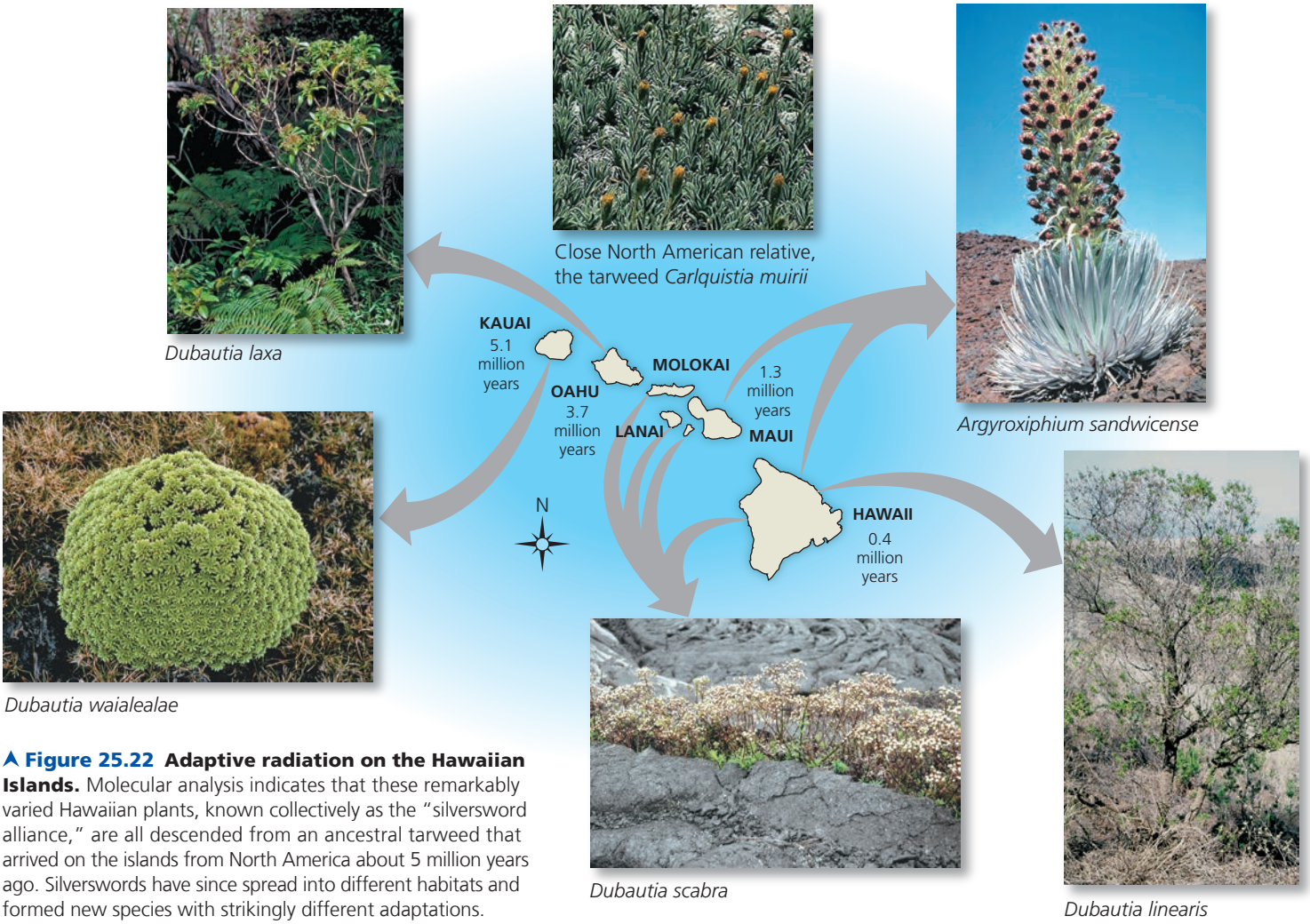


Figure 25.22 Adaptive radiation on the Hawaiian Islands. Molecular analysis indicates that these remarkably varied Hawaiian plants, known collectively as the "silversword alliance," are all descended from an ancestral tarweed that arrived on the islands from North America about 5 million years ago. Silverswords have since spread into different habitats and formed new species with strikingly different adaptations.

island, Hawaii, is less than a million years old and still has active volcanoes. Each island was born "naked" and was gradually populated by stray organisms that rode the ocean currents and winds either from far-distant land areas or from older islands of the archipelago itself. The physical diversity of each island, including immense variation in soil conditions, elevation, and rainfall, provides many opportunities for evolutionary divergence by natural selection. Multiple invasions followed by speciation events have ignited an explosion of adaptive radiation in Hawaii. As a result, thousands of species that inhabit the islands are found nowhere else on Earth. Among plants, for example, about 1,100 species are unique to the Hawaiian Islands. Unfortunately, many of these species are now facing an elevated risk of extinction due to human actions such as habitat destruction and the introduction of non-native plant species.

Animation: Overview of Macroevolution

CONCEPT CHECK 25.4

1. Explain the consequences of plate tectonics for life on Earth.
2. What factors promote adaptive radiations?
3. **WHAT IF?** ➤ Suppose that an invertebrate species was lost in a mass extinction caused by a sudden catastrophic event. Would the last appearance of this species in the fossil record necessarily be close to when the extinction actually occurred? Would the answer to this question differ depending on whether the species was common (abundant and widespread) or rare? Explain.

For suggested answers, see Appendix A.

CONCEPT 25.5

Major changes in body form can result from changes in the sequences and regulation of developmental genes

The fossil record tells us what the great changes in the history of life have been and when they occurred. Moreover, an understanding of plate tectonics, mass extinction, and adaptive radiation provides a picture of how those changes came about. But we can also seek to understand the intrinsic biological mechanisms that underlie changes seen in the fossil record. For this, we turn to genetic mechanisms of change, paying particular attention to genes that influence development.

Effects of Developmental Genes

As you read in Concept 20.6, "evo-devo"—research at the interface between evolutionary biology and developmental biology—is illuminating how slight genetic differences can produce major morphological differences between species. In particular, large morphological differences can result from genes that alter the rate, timing, and spatial pattern of change in an organism's form as it develops from a zygote into an adult.

Changes in Rate and Timing

Many striking evolutionary transformations are the result of **heterochrony** (from the Greek *hetero*, different, and *chronos*, time), an evolutionary change in the rate or timing of developmental events. For example, an organism's shape depends in part on the relative growth rates of different body parts during development. Changes to these rates can alter the adult form substantially, as seen in the contrasting shapes of human and chimpanzee skulls **(Figure 25.23)**. Other examples of the dramatic evolutionary effects of heterochrony include how increased growth rates of finger bones yielded the skeletal structure of wings in bats (see Figure 21.15) and how slowed growth of leg and pelvic bones led to the reduction and eventual loss of hind limbs in whales (see Figure 21.20).

Heterochrony can also alter the timing of reproductive development relative to the development of nonreproductive organs. If the development of reproductive organs accelerates compared to that of other organs, the sexually mature stage of

Figure 25.23 Relative skull growth rates. In the human evolutionary lineage, mutations slowed the growth of the jaw relative to other parts of the skull. As a result, in humans the skull of an adult is more similar to the skull of an infant than is the case for chimpanzees.

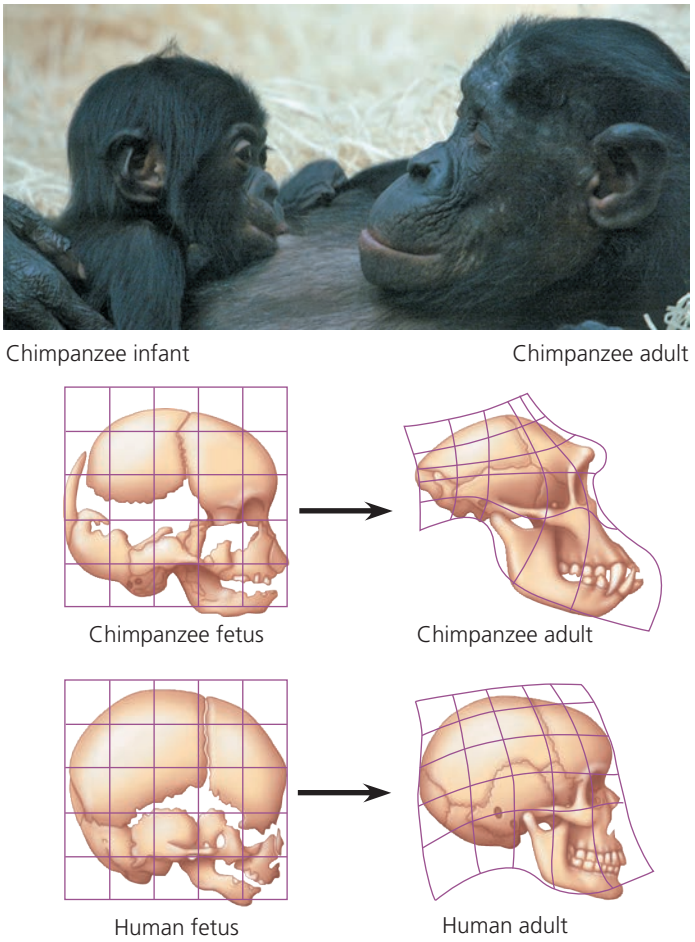


Figure 25.24 Paedomorphosis. The adults of some species retain features that were juvenile in ancestors. This salamander is an axolotl, an aquatic species becomes a sexually mature adult while retaining certain larval (tadpole) characteristics, including gills.

a species may retain body features that were juvenile structures in an ancestral species, a condition called **paedomorphosis** (from the Greek *paedos*, of a child, and *morphosis*, formation). For example, most salamander species have aquatic larvae that undergo metamorphosis in becoming adults. But some species grow to adult size and become sexually mature while retaining gills and other larval features **(Figure 25.24)**. Such an evolutionary alteration of developmental timing can produce animals that appear very different from their ancestors, even though the overall genetic change may be small. Indeed, recent evidence indicates that a change at a single locus was probably sufficient to bring about paedomorphosis in the axolotl salamander, although other genes may have contributed as well.

Changes in Spatial Pattern

Substantial evolutionary changes can also result from alterations in genes that control the spatial organization of body parts. For example, master regulatory genes called **homeotic genes** (see Concept 18.1) determine such basic features as where a pair of wings and legs will develop on a bird or how a plant's flower parts are arranged.

The products of one class of homeotic genes, the *Hox* genes, provide positional information in an animal embryo. This information prompts cells to develop into structures appropriate for a particular location. Changes in *Hox* genes or in how they are expressed can have a profound impact on morphology. For example, among crustaceans, a change in the location where two *Hox* genes (*Ubx* and *Scr*) are expressed correlates with the conversion of a swimming appendage to a feeding appendage. Similarly, when comparing plant species, changes to the expression of homeotic genes known as *MADS-box* genes can produce flowers that differ dramatically in form (see Concept 35.5).

The Evolution of Development

The 560-million-year-old fossils of Ediacaran animals in Figure 25.5 suggest that a set of genes sufficient to produce complex animals existed at least 25 million years *before* the Cambrian explosion. If such genes have existed for so long, how can we explain the astonishing increases in diversity seen during and since the Cambrian explosion?

Adaptive evolution by natural selection provides one answer to this question. As we've seen throughout this unit, by sorting among differences in the sequences of protein-encoding genes, selection can improve adaptations rapidly. In addition, new genes (created by gene duplication events) can take on new metabolic and structural functions, as can existing genes that are regulated in new ways.

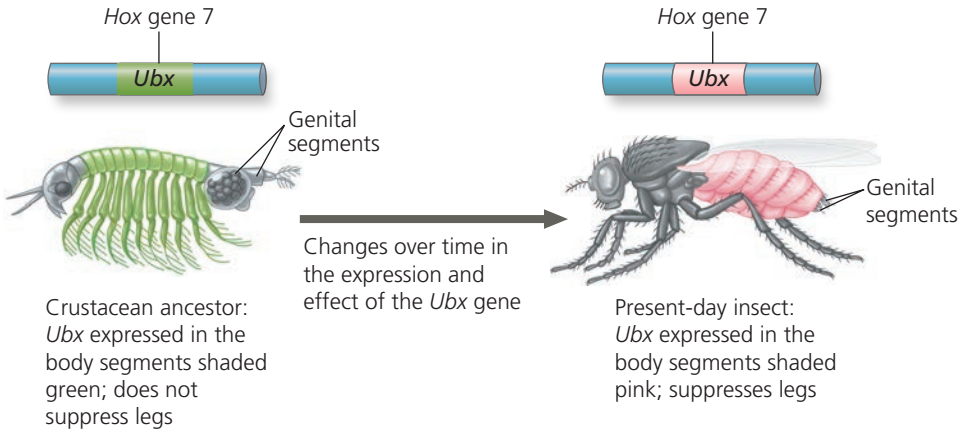
Examples in the previous section suggest that developmental genes may have been particularly important. Thus, we'll turn next to how new morphological forms can arise from changes in the nucleotide sequences or regulation of developmental genes.

Changes in Gene Sequence

New developmental genes arising after gene duplication events probably facilitated the origin of novel morphological forms. But since other genetic changes also may have occurred at such times, it can be difficult to establish causal links between genetic and morphological changes that occurred in the past.

This difficulty was sidestepped in a study of developmental changes associated with the divergence of six-legged insects from crustacean ancestors that had more than six legs. (As discussed in Concept 33.4, insects arose from within a subgroup of the crustaceans, the traditional name for organisms such as shrimp, crabs, and lobsters.) Researchers noted differences between crustaceans and insects in the pattern of expression and the effects of the *Hox* gene *Ubx*: In particular, in insects, *Ubx* suppresses leg formation where it is expressed **(Figure 25.25)**.

Figure 25.25 Effects of the *Hox* gene *Ubx* on the insect body plan. In crustaceans, the *Hox* gene *Ubx* is expressed in the region shaded green, the body segments between the head and genital segments. In insects, *Ubx* is expressed in only a subset (shaded pink) of the homologous body segments, where it suppresses leg formation.



To examine the workings of this gene, researchers cloned the *Ubx* gene from an insect, the fruit fly *Drosophila*, and from a crustacean, the brine shrimp *Artemia*. Next, they genetically engineered fruit fly embryos to express either the *Drosophila Ubx* gene or the *Artemia Ubx* gene throughout their bodies. The *Drosophila* gene suppressed 100% of the limbs in the embryos, as expected, whereas the *Artemia* gene suppressed only 15%.

The researchers then sought to uncover key steps involved in the evolutionary transition from an ancestral *Ubx* gene to an insect *Ubx* gene. Their approach was to identify mutations that would cause the *Artemia Ubx* gene to suppress leg formation, thus making its gene act more like an insect *Ubx* gene. To do this, they constructed a series of "hybrid" *Ubx* genes, each of which contained known segments of the *Drosophila Ubx* gene and known segments of the *Artemia Ubx* gene. By inserting these hybrid genes into fruit fly embryos (one hybrid gene per embryo) and observing their effects on leg development, the researchers were able to pinpoint the exact amino acid changes responsible for the suppression of additional limbs in insects. In so doing, this study provided evidence that particular changes in the nucleotide sequence of a developmental gene contributed to a major evolutionary change: the origin of the six-legged insect body plan.

Changes in Gene Regulation

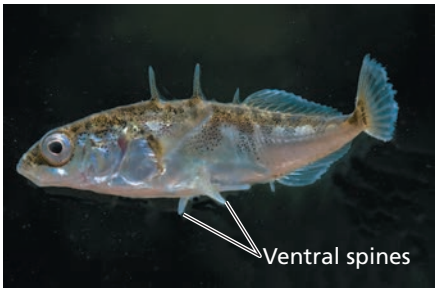
A change in the nucleotide sequence of a gene may affect its function wherever the gene is expressed, while changes in the regulation of gene expression can be limited to one cell type (see Concept 18.4). Thus, a change in the regulation of a developmental gene may have fewer harmful side effects than a change to the sequence of the gene. This reasoning has prompted researchers to suggest that changes in the form of organisms may often be caused by mutations that affect the regulation of developmental genes—not their sequences.

This idea is supported by studies of a variety of species, including threespine stickleback fish. These fish live in the open

Figure 25.26

Inquiry What causes the loss of spines in lake stickleback fish?

Experiment Marine populations of the threespine stickleback fish (*Gasterosteus aculeatus*) have a set of protective spines on their lower (ventral) surface; however, these spines have been lost or reduced in some lake populations of this fish. Working at Stanford University, Michael Shapiro, David Kingsley, and colleagues performed genetic crosses and found that most of the reduction in spine size resulted from the effects of a single developmental gene, *Pitx1*. The researchers then tested two hypotheses about how *Pitx1* causes this morphological change.



▲ **Threespine stickleback (*Gasterosteus aculeatus*)**

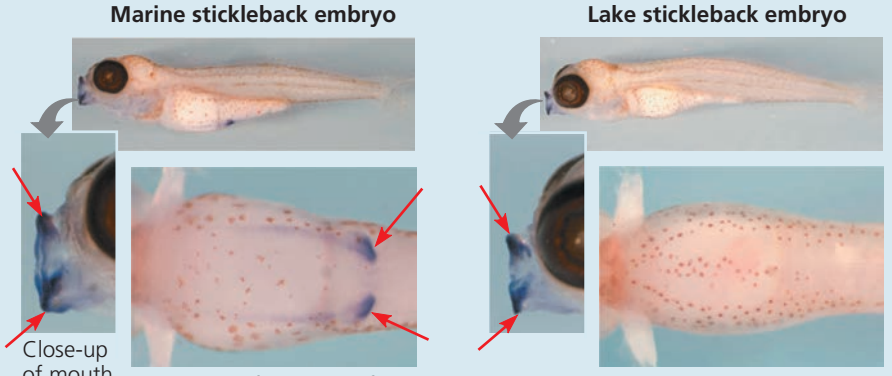
Hypothesis A: A change in the DNA sequence of *Pitx1* had caused spine reduction in lake populations. To test this idea, the team used DNA sequencing to compare the coding sequence of the *Pitx1* gene between marine and lake stickleback populations.

Hypothesis B: A change in the regulation of the expression of *Pitx1* had caused spine reduction. To test this idea, the researchers monitored where in the developing embryo the *Pitx1* gene was expressed. They conducted whole-body *in situ* hybridization experiments (see Concept 19.2) using *Pitx1* DNA as a probe to detect *Pitx1* mRNA in the fish.

Results

Test of Hypothesis A: Are there differences in the coding sequence of the *Pitx1* gene in marine and lake stickleback fish? → **Result: No** The 283 amino acids of the *Pitx1* protein are identical in marine and lake stickleback populations.

Test of Hypothesis B: Are there any differences in the regulation of expression of *Pitx1?* → **Result: Yes** Red arrows (→) indicate regions of *Pitx1* gene expression in the photographs below. *Pitx1* is expressed in the ventral spine and mouth regions of developing marine stickleback fish but only in the mouth region of developing lake stickleback fish.



Conclusion The loss or reduction of ventral spines in lake populations of threespine stickleback fish appears to have resulted primarily from a change in the regulation of *Pitx1* gene expression, not from a change in the gene's sequence.

Data from M. D. Shapiro et al., Genetic and developmental basis of evolutionary pelvic reduction in three-spine sticklebacks, *Nature* 428:717–723 (2004).

WHAT IF? ➤ *Describe the set of results that would have led researchers to the conclusion that a change in the coding sequence of the* Pitx1 *gene was more important than a change in regulation of gene expression.*

ocean and in shallow, coastal waters. In western Canada, they also live in lakes formed when the coastline receded during the past 12,000 years. Marine stickleback fish have a pair of spines on their ventral (lower) surface, which deter some predators. These spines are often reduced or absent in stickleback fish living in lakes that lack predatory fishes and that are also low in calcium. Spines may have been lost in such lakes because they are not advantageous in the absence of predators, and the limited calcium is needed for purposes other than constructing spines.

At the genetic level, the developmental gene *Pitx1* was known to influence whether stickleback fish have ventral spines. Was the reduction of spines in some lake populations due to changes in the *Pitx1* gene or to changes in how the gene is expressed **(Figure 25.26)**? The researchers' results indicate that the regulation of gene expression has changed, not the DNA sequence. Moreover, lake stickleback fish do express the *Pitx1* gene in tissues not related to the production of spines (for example, the mouth), illustrating how morphological change can be caused by altering the expression of a developmental gene in some parts of the body but not others. In a 2010 follow-up study, researchers showed that changes to the *Pel* enhancer, a noncoding DNA region that affects expression of the *Pitx1* gene, resulted in the reduction of ventral spines in lake sticklebacks. Overall, results from studies on stickleback fish provide a clear and detailed example of how changes in gene regulation can alter the form of individual organisms and ultimately lead to evolutionary change in populations.

HHMI Video: The Making of the Fittest: Evolving Switches, Evolving Bodies (Stickleback)

CONCEPT CHECK 25.5

1. Explain how new body forms can originate by heterochrony.
2. Why is it likely that *Hox* genes have played a major role in the evolution of novel morphological forms?
3. **MAKE CONNECTIONS ➤** Given that changes in morphology are often caused by changes in the regulation of gene expression, predict whether noncoding DNA is likely to be affected by natural selection. See Concept 18.3 to review noncoding DNA and regulation of gene expression.

For suggested answers, see Appendix A.

CONCEPT 25.6

Evolution is not goal oriented

What does our study of macroevolution tell us about how evolution works? One lesson is that throughout the history of life, the origin of new species has been affected by both the small-scale factors described in Concept 23.3 (such as natural selection operating in populations) and the large-scale factors described in this chapter (such as continental drift promoting bursts of speciation throughout the globe). Moreover, to paraphrase the Nobel Prize–winning geneticist François Jacob, evolution is like tinkering—a process in which new forms arise by the modification of existing structures or existing developmental genes. Over time, such tinkering has led to the three key features of the natural world described on the opening page of Chapter 21: the striking ways in which organisms are suited for life in their environments, the many shared characteristics of life, and the rich diversity of life.

Evolutionary Novelties

François Jacob's view of evolution harkens back to Darwin's concept of descent with modification. As new species form, novel and complex structures can arise as gradual modifications of ancestral structures. In many cases, complex structures have evolved in increments from simpler versions that performed the same basic function. For example, consider the human eye, an intricate organ constructed from numerous parts that work together in forming an image and transmitting it to the brain. How could the human eye have evolved in gradual increments? Some argue that if the eye needs all of its components to function, a partial eye could not have been of use to our ancestors.

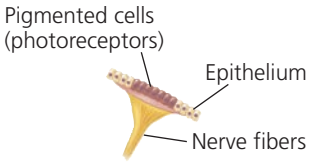
The flaw in this argument, as Darwin himself noted, lies in the assumption that only complicated eyes are useful. In fact, many animals depend on eyes that are far less complex than our own. The simplest eyes that we know of are patches of light-sensitive photoreceptor cells. These simple eyes appear to have had a single evolutionary origin and are now found in a variety of animals, including small molluscs called limpets. Such eyes have no equipment for focusing images, but they do enable the animal to distinguish light from dark. Limpets cling more tightly to their rock when a shadow falls on them, a behavioral adaptation that reduces the risk of being eaten **(Figure 25.27)**. Limpets have had a long evolutionary history, demonstrating that their "simple" eyes are quite adequate to support their survival and reproduction.

In the animal kingdom, complex eyes have evolved independently from such basic structures many times. Some molluscs, such as squids and octopuses, have eyes as complex as those

Figure 25.27 Limpets (*Patella vulgata*), molluscs that can sense light and dark with a simple patch of photoreceptor cells.

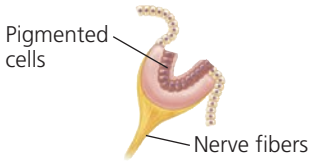
Figure 25.28 A range of eye complexity among molluscs.

(a) Patch of pigmented cells



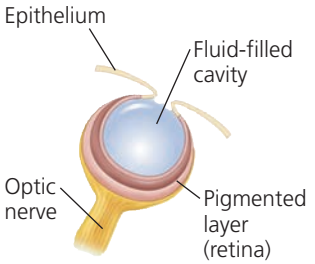
The limpet *Patella* has a simple patch of photoreceptors.

(b) Eyecup



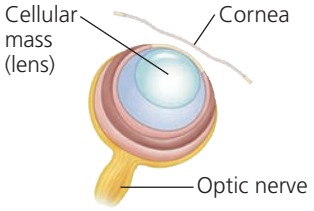
The slit shell mollusc *Pleurotomaria* has an eyecup.

(c) Pinhole camera-type eye



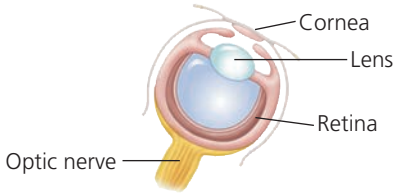
The *Nautilus* eye functions like a pinhole camera (an early type of camera lacking a lens).

(d) Eye with primitive lens



The marine snail *Murex* has a primitive lens consisting of a mass of crystal-like cells. The cornea is a transparent region of tissue that protects the eye and helps focus light.

(e) Complex camera lens-type eye



The squid *Loligo* has a complex eye with features (cornea, lens, and retina) similar to those of vertebrate eyes. However, the squid eye evolved independently from vertebrate eyes.

of humans and other vertebrates **(Figure 25.28)**. Although complex mollusc eyes evolved independently of vertebrate eyes, both evolved from a simple cluster of photoreceptor cells present in a common ancestor. In each case, the complex eye evolved through a series of steps that benefited the eyes' owners at every stage. Evidence of their independent evolution can be found in their structure: Vertebrate eyes detect light at the back layer of the retina and conduct nerve impulses toward the front, while complex mollusc eyes do the reverse.

Throughout their evolutionary history, eyes retained their basic function of vision. But evolutionary novelties can also arise when structures that originally played one role gradually acquire a different one. For example, as cynodonts gave rise to early mammals, bones that formerly comprised the jaw hinge (the articular and quadrate; see Figure 25.7) were incorporated into the ear region of mammals, where they eventually took on a new function: the transmission of sound (see Concept 34.6). Structures that evolve in one context but become co-opted for another function are sometimes called *exaptations* to distinguish them from the adaptive origin of the original structure. Note that the concept of exaptation does not imply that a structure somehow evolves in anticipation of future use. Natural selection cannot predict the future; it can only improve a structure in the context of its *current* utility. Novel features, such as the new jaw hinge and ear bones of early mammals, can arise gradually via a series of intermediate stages, each of which has some function in the organism's current context.

Evolutionary Trends

What else can we learn from patterns of macroevolution? Consider evolutionary "trends" observed in the fossil record. For instance, some evolutionary lineages exhibit a trend toward larger or smaller body size. An example is the evolution of the present-day horse (genus *Equus*), a descendant of the 55-million-year-old *Hyracotherium* **(Figure 25.29)**. About the size of a large dog, *Hyracotherium* had four toes on its front feet, three toes on its hind feet, and teeth adapted for browsing on bushes and trees. In comparison, present-day horses are larger, have only one toe on each foot, and possess teeth modified for grazing on grasses.

Extracting a single evolutionary progression from the fossil record can be misleading, however; it is like describing a bush as growing toward a single point by tracing only the branches that lead to that twig. For example, by selecting certain species from the available fossils, it is possible to arrange a succession of animals intermediate between *Hyracotherium* and living horses that shows a trend toward large, single-toed species (follow the yellow highlighting in Figure 25.29). However, if we consider *all* fossil horses known today, this apparent trend vanishes. The genus *Equus* did not evolve in a straight line; it is the only surviving twig of an evolutionary tree that is so branched that it is more like a bush. *Equus* actually descended through a series of speciation episodes that included several adaptive radiations, not all of which led to large, one-toed, grazing horses. In fact, phylogenetic analyses suggest that all lineages that include grazers are closely related to *Parahippus*; the many other horse lineages, all of which are now extinct, remained multi-toed browsers for 35 million years.

Branching evolution *can* result in a real evolutionary trend even if some species counter the trend. One model of long-term trends views species as analogous to individuals: Speciation is their birth, extinction is their death, and new species that diverge from them are their offspring. In this model, just as populations of individual organisms undergo natural selection, species undergo *species selection*. The species that endure the longest and generate the most new offspring species determine the direction of major evolutionary trends. The species selection model suggests that "differential speciation success" plays a role in macroevolution similar to the role of differential reproductive success in microevolution. Evolutionary trends can

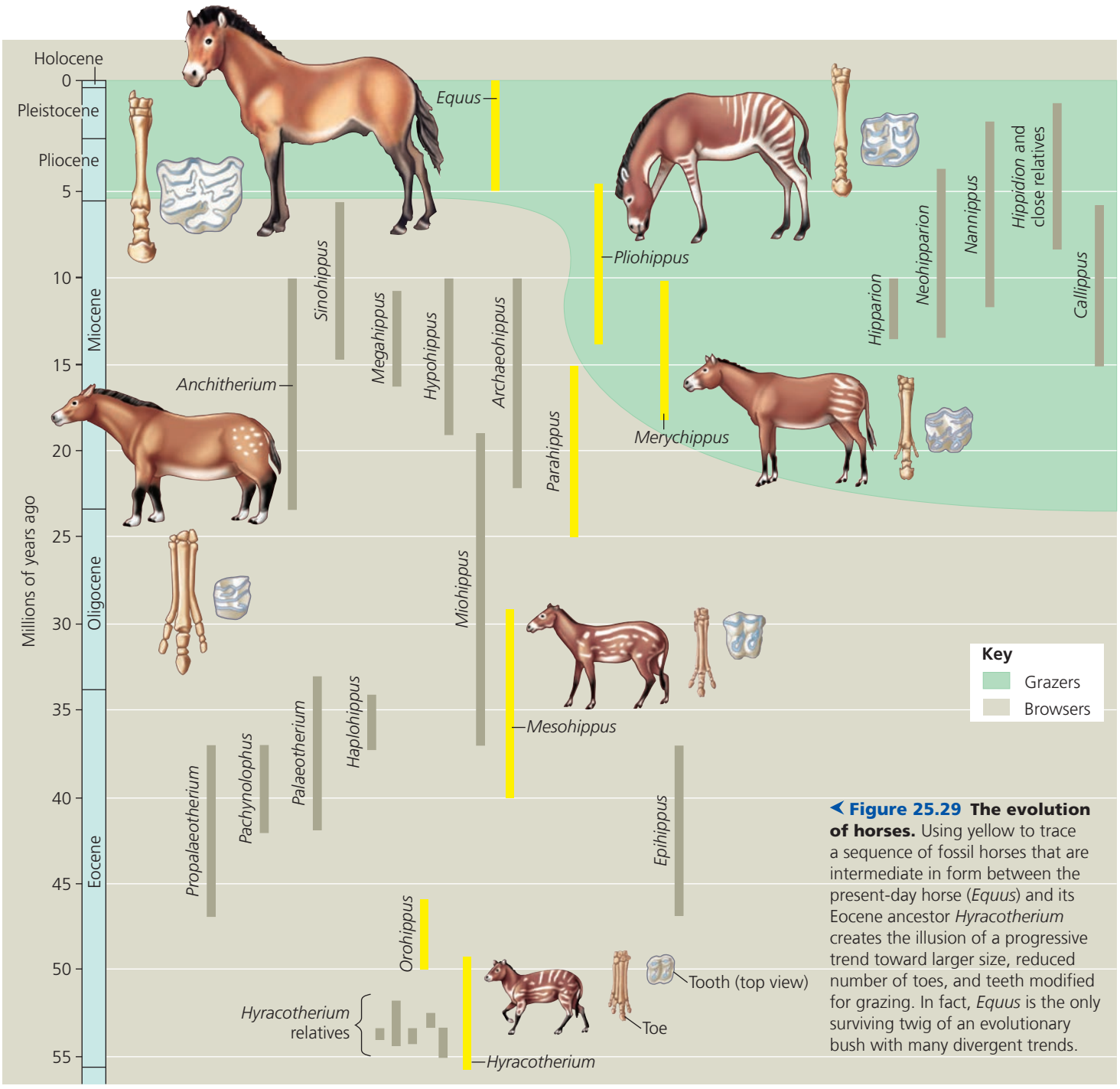


Figure 25.29 The evolution of horses. Using yellow to trace a sequence of fossil horses that are intermediate in form between the present-day horse (*Equus*) and its Eocene ancestor *Hyracotherium* creates the illusion of a progressive trend toward larger size, reduced number of toes, and teeth modified for grazing. In fact, *Equus* is the only surviving twig of an evolutionary bush with many divergent trends.

also result directly from natural selection. For example, when horse ancestors invaded the grasslands that spread during the mid-Cenozoic, there was strong selection for grazers that could escape predators by running faster. This trend would not have occurred without open grasslands.

Whatever its cause, an evolutionary trend does not imply that there is some intrinsic drive toward a particular phenotype. Evolution is the result of the interactions between organisms and their current environments; if environmental conditions change, an evolutionary trend may cease or even reverse itself. The cumulative effect of these ongoing interactions between organisms and their environments is enormous: It is through them that the staggering diversity of life—Darwin's "endless forms most beautiful"—has arisen.

CONCEPT CHECK 25.6

1. How can the Darwinian concept of descent with modification explain the evolution of such complex structures as the vertebrate eye?
2. **WHAT IF?** The myxoma virus kills up to 99.8% of infected European rabbits in populations with no previous exposure to the virus. The virus is transmitted between living rabbits by mosquitoes. Describe an evolutionary trend (in either the rabbit or virus) that might occur after a rabbit population first encounters the virus.

For suggested answers, see Appendix A.

25 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 25.1

Conditions on early Earth made the origin of life possible *(pp. 580–582)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

- Experiments simulating possible early atmospheres have produced organic molecules from inorganic precursors. Amino acids, lipids, sugars, and nitrogenous bases have also been found in meteorites.
- Amino acids and RNA nucleotides polymerize when dripped onto hot sand, clay, or rock. Organic compounds can spontaneously assemble into **protocells**, membrane-bounded droplets that have some properties of cells.
- The first genetic material may have self-replicating, catalytic RNA. Early protocells containing such RNA would have increased through natural selection.

? *Describe the roles that montmorillonite clay and vesicles may have played in the origin of life.*

CONCEPT 25.2

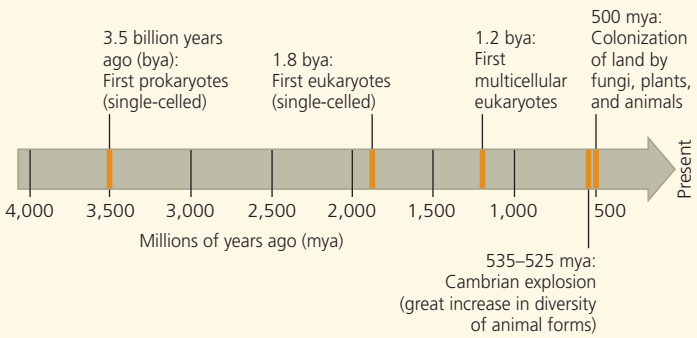
The fossil record documents the history of life *(pp. 582–586)*

- The fossil record, based largely on fossils found in sedimentary rocks, documents the rise and fall of different groups of organisms over time.
- Sedimentary strata reveal the relative ages of fossils. The ages of fossils can be estimated by **radiometric dating** and other methods.
- The fossil record shows how new groups of organisms can arise via the gradual modification of preexisting organisms.

? *What are the challenges of estimating the ages of old fossils? Explain how these challenges may be overcome in some circumstances.*

CONCEPT 25.3

Key events in life's history include the origins of unicellular and multicellular organisms and the colonization of land *(pp. 586–591)*



? *What is the "Cambrian explosion," and why is it significant?*

CONCEPT 25.4

The rise and fall of groups of organisms reflect differences in speciation and extinction rates *(pp. 591–598)*

- In **plate tectonics**, continental plates move gradually over time, altering the physical geography and climate of Earth, leading to extinctions in some groups and speciation in others.
- Evolutionary history has been punctuated by five **mass extinctions** that radically altered life's history. Possible causes for these extinctions include continental drift, volcanic activity, and impacts from comets.
- Large increases in the diversity of life have resulted from **adaptive radiations** that followed mass extinctions. Adaptive radiations have also occurred in groups of organisms that possessed major evolutionary innovations or that colonized new regions in which there was little competition from other organisms.

? *Explain how the broad evolutionary changes seen in the fossil record are the cumulative result of speciation and extinction events.*

CONCEPT 25.5

Major changes in body form can result from changes in the sequences and regulation of developmental genes *(pp. 598–601)*

- Developmental genes affect morphological differences between species by influencing the rate, timing, and spatial patterns of change in an organism's form as it develops into an adult.
- The evolution of new forms can be caused by changes in the nucleotide sequences or regulation of developmental genes.

? *How could changes in a single gene or DNA region ultimately lead to the origin of a new group of organisms?*

CONCEPT 25.6

Evolution is not goal oriented *(pp. 601–603)*

- Novel and complex biological structures can evolve through a series of incremental modifications, each of which benefits the organism that possesses it.
- Evolutionary trends can be caused by natural selection in a changing environment or species selection, resulting from interactions between organisms and their current environments.

? *Explain the reasoning behind the statement "Evolution is not goal oriented."*

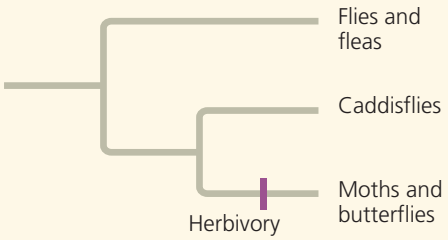
TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–7 can be found in the Study Area in MasteringBiology.

8. **EVOLUTION CONNECTION** Describe how gene flow, genetic drift, and natural selection all can influence macroevolution.

PRACTICE TEST goo.gl/iAsVgL

9. **SCIENTIFIC INQUIRY** Herbivory (plant eating) has evolved repeatedly in insects, typically from meat-eating or detritus-feeding ancestors (detritus is dead organic matter). Moths and butterflies, for example, eat plants, whereas their "sister group" (the insect group to which they are most closely related), the caddisflies, feed on animals, fungi, or detritus. As illustrated in the following phylogenetic tree, the combined moth/butterfly and caddisfly group shares a common ancestor with flies and fleas. Like caddisflies, flies and fleas are thought to have evolved from ancestors that did not eat plants.



There are 140,000 species of moths and butterflies and 7,000 species of caddisflies. State a hypothesis about the impact of herbivory on adaptive radiations in insects. How could this hypothesis be tested?

10. **WRITE ABOUT A THEME: ORGANIZATION** You have seen many examples of how form fits function at all levels of the biological hierarchy. However, we can imagine forms that would function better than some forms actually found in nature. For example, if the wings of a bird were not formed from its forelimbs, such a hypothetical bird could fly yet also hold objects with its forelimbs. In a short essay (100–150 words), use the concept of "evolution as tinkering" to explain why there are limits to the functionality of forms in nature.

11. **SYNTHESIZE YOUR KNOWLEDGE**

In 2010, the Soufriere Hills volcano on the Caribbean island of Montserrat erupted violently, spewing huge clouds of ash and gases into the sky. Explain how the volcanic eruptions at the end of the Permian period and the formation of Pangaea, both of which occurred about 252 million years ago, set in motion events that altered evolutionary history.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

This page is intentionally left blank.

UNIT 5 THE DIVERSITY OF LIFE

Winner of a MacArthur Fellowship "Genius" Award, Nancy A. Moran is the Leslie Surginer Endowed Professor of Integrative Biology at the University of Texas at Austin. Dr. Moran received a B.A. in general studies from the University of Texas, graduating with highest honors, and a Ph.D. in zoology from the University of Michigan. Elected to the U.S. National Academy of Sciences in 2004, Dr. Moran has a long-standing interest in the evolutionary history and ecological importance of symbioses. Currently, she and her students are using experimental and phylogenomic approaches to study beneficial interactions between insects such as aphids and honeybees and bacteria that live within their bodies.

An Interview with Nancy Moran

How did you first become interested in science?

As a kid, I liked insects and enjoyed gardening. However, I did not see myself as becoming a scientist, and I began college as an art major. But for me, art was just too hard, and I switched into an honors program that did not require you to declare a major. As a junior, I happened to take a class in biology, which I really enjoyed. So I took more biology courses and eventually did an honors thesis on mate choice in birds. By then, I was hooked!

Much of your work is on symbioses. What are symbioses, and what drew you to this topic?

A symbiosis is a close physiological relationship in which one species (the symbiont) lives in or on a larger species (the host). In the late 1980s, I was working on aphids, which have many bacterial symbionts. No one knew what these symbionts did because they can't live outside of their host. I wasn't planning to study symbioses, but then I got an intriguing phone call from the bacteriologist Paul Baumann. At that time, PCR had just become widely available, which enabled researchers to amplify and sequence particular genes. Paul had the foresight to realize that we could use PCR to ask, "What kinds of bacteria live within aphids as symbionts, and based on their genes, what are these symbionts doing?"

"We sequenced these genes and determined that they came from a fungus. What a surprise—the carotenoid genes of a fungus had become part of the DNA of an aphid!"

How do symbioses affect life on Earth?

Symbioses are everywhere. Every organism is engaged in symbioses, including humans. As one example, our bodies contain bacteria that break down carbohydrates that we otherwise could not digest. In the case of aphids, these insects feed on plant sap, which provides sugars but lacks other essential nutrients, such as key amino acids. These amino acids are supplied to the aphids by bacterial symbionts. At a higher level, we can say that entire ecosystems depend on symbioses. Coral reefs are built by coral polyps that depend on energy provided by algal symbionts. Without their symbionts, the corals can die, harming the reef and the many fishes and other organisms that live there.

Adult female pea aphids and asexually produced daughters. The reddish color is due to carotenoid pigments.

What is the most surprising discovery that you've made?

One such discovery concerns carotenoids, which are colored molecules used in photosynthesis by plants and in light detection by many animals. It was thought that animals could not make carotenoids and so had to get them from their diet. While studying the genome of an aphid, we were startled to find that it had functional carotenoid-synthesizing genes. We sequenced these genes and determined that they came from a fungus. What a surprise—the carotenoid genes of a fungus had become part of the DNA of an aphid!

What advice do you have for students considering a career in biology?

Every student is different. Some students categorize themselves as not interested in or unable to "do" science. But the only way to tell is to try it: Dive in, see if you enjoy research, and discover the special skills that you bring to the table. Biology needs people with different skills. One person might excel at communication, another is meticulous, while a third can synthesize many different pieces of information. If you find that you enjoy biology, go for it—whether you are a freshman or a senior, it is not too late.

26 Introduction to Viruses

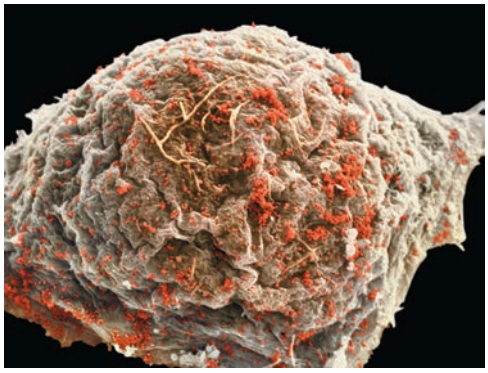
Figure 26.1 Are the viruses (reddish-purple) budding from these cells alive?

KEY CONCEPTS

26.1 A virus consists of a nucleic acid surrounded by a protein coat

26.2 Viruses replicate only in host cells

26.3 Viruses and prions are formidable pathogens in animals and plants



A human immune cell infected with HIV. New viruses (red) are budding from the plasma membrane (colorized SEM).

A Borrowed Life

The illustration in **Figure 26.1** shows a remarkable event: Human immune cells (purple) infected by human immunodeficiency viruses (HIV) are releasing more HIV viruses. These viruses (red, surrounded by a protein-studded purple membrane from the immune cell) will infect other cells. (The SEM below shows one infected cell.) By injecting its genetic information into a cell, a virus hijacks the cell, recruiting cellular machinery to manufacture many new viruses and promote further infection. Left untreated, HIV causes acquired immunodeficiency syndrome (AIDS) by destroying vital immune system cells.

Compared with eukaryotic and even prokaryotic cells, viruses are much smaller and simpler in structure. Lacking the structures and metabolic machinery found in a cell, a **virus** is an infectious particle consisting of little more than genes packaged in a protein coat.

Are viruses living or nonliving? Early on, they were considered biological chemicals; the Latin root for virus means "poison." Viruses are capable of causing a wide variety of diseases, so researchers in the late 1800s saw a parallel with bacteria and proposed that viruses were the simplest of living forms. However, viruses cannot reproduce or carry out metabolic activities outside of a host cell. Most biologists studying viruses today would probably agree that they are not alive but exist in a shady area between life-forms and chemicals. The simple phrase used recently by two researchers describes them aptly enough: Viruses lead "a kind of borrowed life."

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

To a large extent, molecular biology was born in the laboratories of biologists studying viruses that infect bacteria. Experiments with these viruses provided evidence that genes are made of nucleic acids, and they were critical in working out the molecular mechanisms of the fundamental processes of DNA replication, transcription, and translation.

In this chapter, we will explore the biology of viruses, beginning with their structure and then describing how they replicate. Next, we will discuss the role of viruses as disease-causing agents, or pathogens, and conclude by considering some even simpler infectious agents called prions.

CONCEPT 26.1

A virus consists of a nucleic acid surrounded by a protein coat

Scientists were able to detect viruses indirectly long before they were actually able to see them. The story of how viruses were discovered begins near the end of the 19th century.

The Discovery of Viruses: *Scientific Inquiry*

Tobacco mosaic disease stunts the growth of tobacco plants and gives their leaves a mottled, or mosaic, coloration. In 1883, Adolf Mayer, a German scientist, discovered that he could transmit the disease from plant to plant by rubbing sap extracted from diseased leaves onto healthy plants. After an unsuccessful search for an infectious microbe in the sap, Mayer suggested that the disease was caused by unusually small bacteria that were invisible under a microscope. This hypothesis was tested a decade later by Dmitri Ivanowsky, a Russian biologist who passed sap from infected tobacco leaves through a filter designed to remove bacteria. After filtration, the sap still produced mosaic disease.

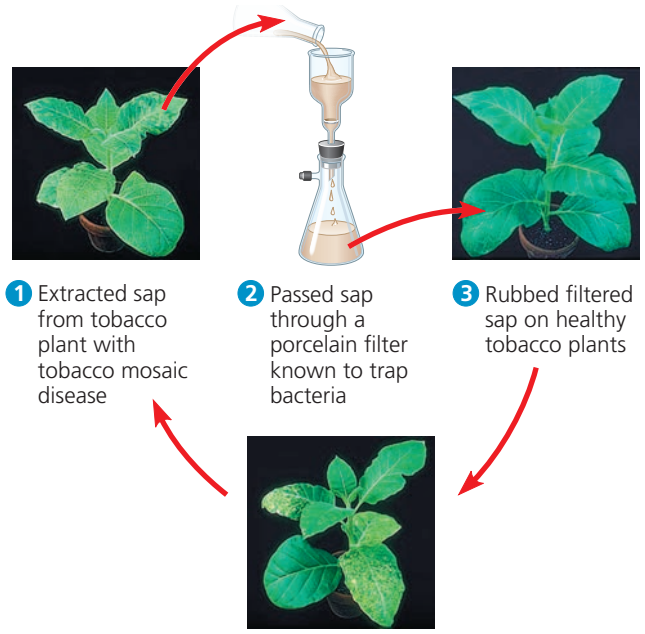
But Ivanowsky clung to the hypothesis that bacteria caused tobacco mosaic disease. Perhaps, he reasoned, the bacteria were small enough to pass through the filter or made a toxin that could do so. The second possibility was ruled out when the Dutch botanist Martinus Beijerinck carried out a classic series of experiments that showed that the infectious agent in the filtered sap could replicate **(Figure 26.2)**.

In fact, the pathogen replicated only within the host it infected. In further experiments, Beijerinck showed that unlike bacteria used in the lab at that time, the mysterious agent of mosaic disease could not be cultivated on nutrient media in test tubes or petri dishes. Beijerinck imagined a replicating particle much smaller and simpler than a bacterium, and he is generally credited with being the first scientist to voice the concept of a virus. His suspicions were confirmed in 1935 when the American scientist Wendell Stanley crystallized the infectious particle, now known as tobacco mosaic virus (TMV). Subsequently, TMV and many other viruses were actually seen with the help of the electron microscope.

Figure 26.2

Inquiry What causes tobacco mosaic disease?

Experiment In the late 1800s, Martinus Beijerinck, of the Technical School in Delft, the Netherlands, investigated the properties of the agent that causes tobacco mosaic disease (then called spot disease).



Results When the filtered sap was rubbed on healthy plants, they became infected. Their sap, extracted and filtered, could then act as a source of infection for another group of plants. Each successive group of plants developed the disease to the same extent as earlier groups.

Conclusion The infectious agent was apparently not a bacterium because it could pass through a bacterium-trapping filter. The pathogen must have been replicating in the plants because its ability to cause disease was undiluted after several transfers from plant to plant.

Data from M. J. Beijerinck, Concerning a *contagium vivum fluidum* as cause of the spot disease of tobacco leaves, *Verhandelingen der Koninkyke akademie Wettenschappen te Amsterdam* 65:3–21 (1898). Translation published in English as Phytopathological Classics Number 7 (1942), American Phytopathological Society Press, St. Paul, MN.

WHAT IF? ➤ *If Beijerinck had observed that the infection of each group was weaker than that of the previous group and that ultimately the sap could no longer cause disease, what might he have concluded?*

Structure of Viruses

The tiniest viruses are only 20 nm in diameter—smaller than a ribosome. Millions could easily fit on a pinhead. Even the largest known virus, which has a diameter of 1,500 nanometers (1.5 μm), is barely visible under the light microscope. Stanley's discovery that some viruses could be crystallized was exciting and puzzling news. Not even the simplest of cells can aggregate into regular crystals. But if viruses are not cells, then what are they? Examining the structure of a virus more closely reveals that it is an infectious particle consisting of nucleic acid enclosed in a protein coat and, for some viruses, surrounded by a membranous envelope.

Viral Genomes

We usually think of genes as being made of double-stranded DNA, but many viruses defy this convention. Their genomes may consist of double-stranded DNA, single-stranded DNA, double-stranded RNA, or single-stranded RNA, depending on the type of virus. A virus is called a DNA virus or an RNA virus based on the kind of nucleic acid that makes up its genome. In either case, the genome is usually organized as a single linear or circular molecule of nucleic acid, although the genomes of some viruses consist of multiple molecules of nucleic acid. The smallest viruses known have only three genes in their genome, while the largest have several hundred to 2,000. For comparison, bacterial genomes contain about 200 to a few thousand genes.

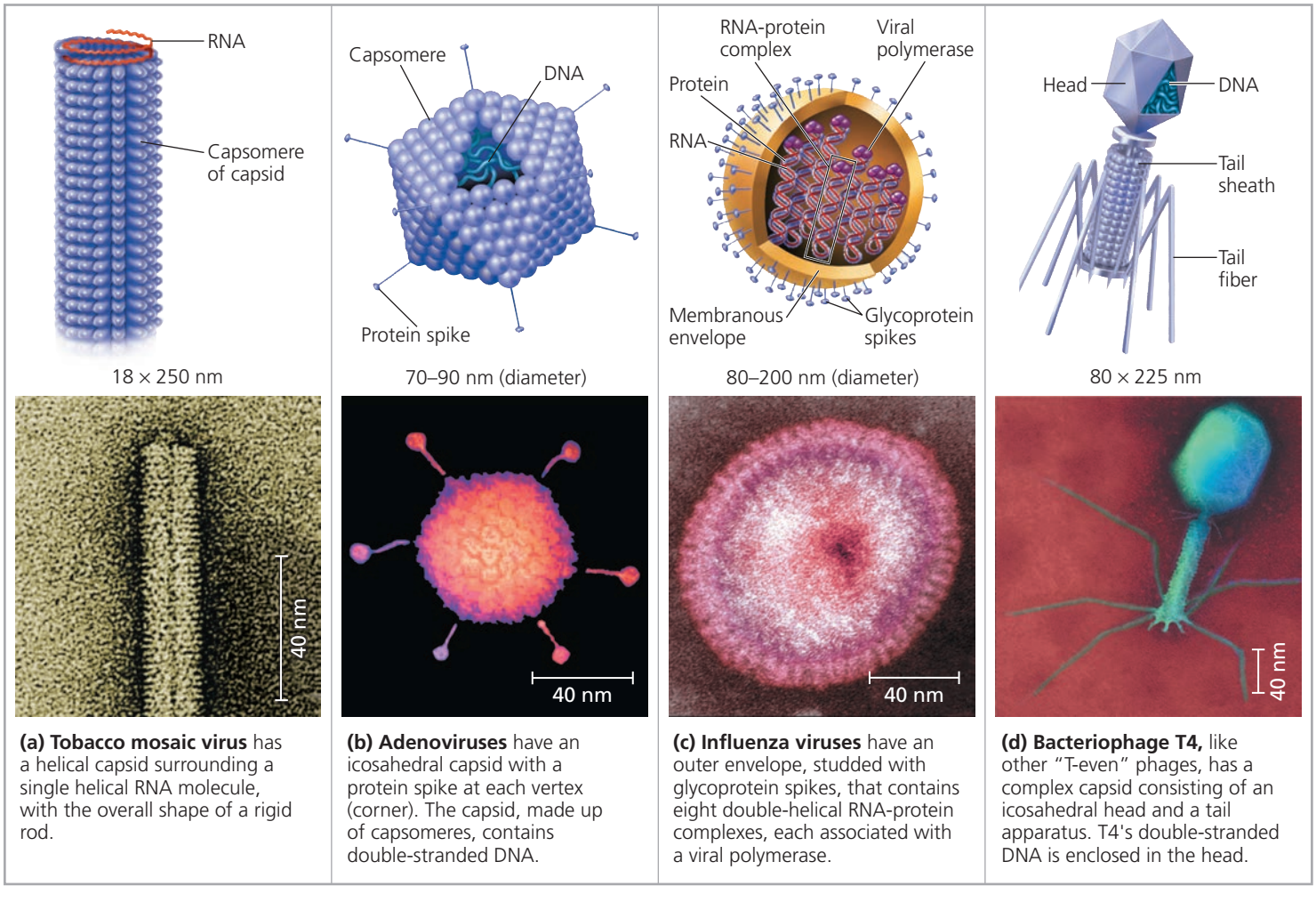
Capsids and Envelopes

The protein shell enclosing the viral genome is called a **capsid**. Depending on the type of virus, the capsid may be rod-shaped, polyhedral, or more complex in shape. Capsids are built from a large number of protein subunits called *capsomeres*, but the number of different *kinds* of proteins in a capsid is usually small. Tobacco mosaic virus has a rigid, rod-shaped capsid made from over 1,000 molecules of a single type of protein arranged in a helix; rod-shaped viruses are commonly called *helical viruses* for this reason **(Figure 26.3a)**. Adenoviruses, which infect the respiratory tracts of animals, have 252 identical protein molecules arranged in a polyhedral capsid with 20 triangular facets—an icosahedron; thus, these and other similarly shaped viruses are referred to as *icosahedral viruses* **(Figure 26.3b)**.

Some viruses have accessory structures that help them infect their hosts. For instance, a membranous envelope surrounds the capsids of influenza viruses and many other viruses found in animals **(Figure 26.3c)**. These **viral envelopes**, which are derived from the membranes of the host cell, contain host cell phospholipids and membrane proteins. They also contain proteins and glycoproteins of viral origin. (Glycoproteins are proteins with carbohydrates covalently attached.) Some viruses carry a few viral enzyme molecules within their capsids.

Many of the most complex capsids are found among the viruses that infect bacteria, called **bacteriophages**,

Figure 26.3 Viral structure. Viruses are made up of nucleic acid (DNA or RNA) enclosed in a protein coat (the capsid) and sometimes further wrapped in a membranous envelope. The individual protein subunits making up the capsid are called capsomeres. Although diverse in size and shape, viruses have many common structural features. (All micrographs are colorized TEMs.)



(a) Tobacco mosaic virus has a helical capsid surrounding a single helical RNA molecule, with the overall shape of a rigid rod.

(b) Adenoviruses have an icosahedral capsid with a protein spike at each vertex (corner). The capsid, made up of capsomeres, contains double-stranded DNA.

(c) Influenza viruses have an outer envelope, studded with glycoprotein spikes, that contains eight double-helical RNA-protein complexes, each associated with a viral polymerase.

(d) Bacteriophage T4, like other "T-even" phages, has a complex capsid consisting of an icosahedral head and a tail apparatus. T4's double-stranded DNA is enclosed in the head.

or simply **phages**. The first phages studied included seven that infect *Escherichia coli*. These seven phages were named type 1 (T1), type 2 (T2), and so forth, in the order of their discovery. The three "T-even" phages (T2, T4, and T6) turned out to be very similar in structure. Their capsids have elongated icosahedral heads enclosing their DNA. Attached to the head is a protein tail piece with fibers by which the phages attach to a bacterial cell **(Figure 26.3d)**. In the next section, we'll examine how these few viral parts function together with cellular components to produce large numbers of viral progeny.

CONCEPT CHECK 26.1

1. **VISUAL SKILLS ➤** Compare the structures of tobacco mosaic virus (TMV) and influenza virus (see Figure 26.3).
2. **MAKE CONNECTIONS ➤** Bacteriophages were used to provide evidence that DNA carries genetic information (see Figure 16.4). Briefly describe the experiment carried out by Hershey and Chase, including in your description why the researchers chose to use phages.

For suggested answers, see Appendix A.

CONCEPT 26.2

Viruses replicate only in host cells

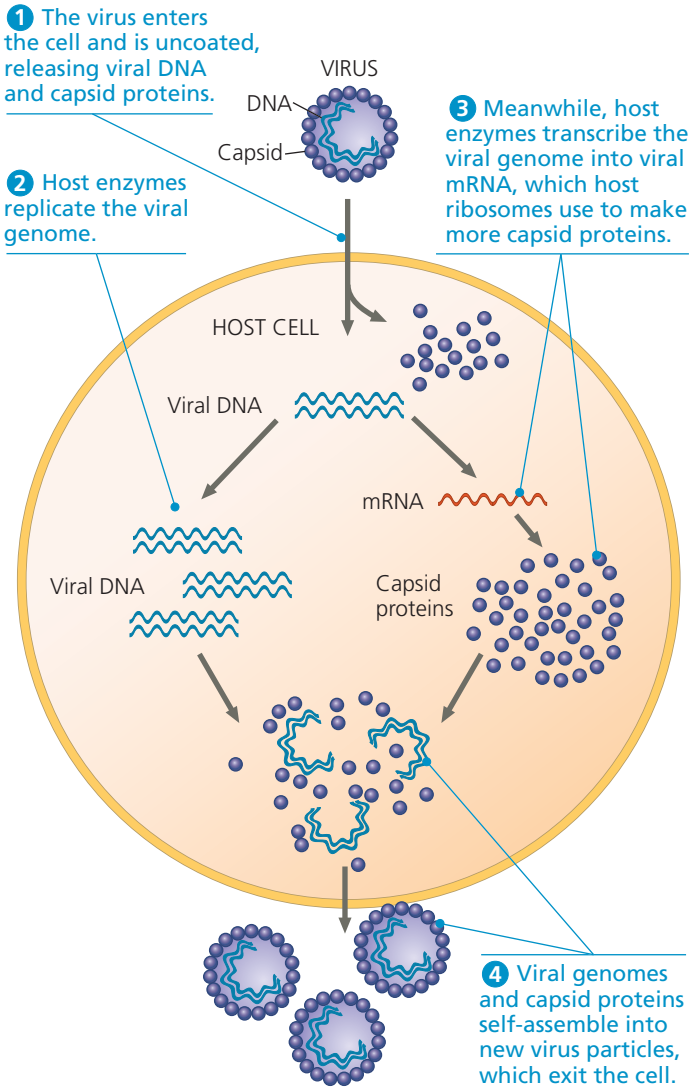
Viruses lack metabolic enzymes and equipment for making proteins, such as ribosomes. They are obligate intracellular parasites; in other words, they can replicate only within a host cell. It is fair to say that viruses in isolation are merely packaged sets of genes in transit from one host cell to another.

Each particular virus can infect cells of only a limited number of host species, called the **host range** of the virus. This host specificity results from the evolution of recognition systems by the virus. Viruses usually identify host cells by a "lock-and-key" fit between viral surface proteins and specific receptor molecules on the outside of cells. According to one model, such receptor molecules originally carried out functions that benefited the host cell but were co-opted later by viruses as portals of entry. Some viruses have broad host ranges. For example, West Nile virus and equine encephalitis virus are distinctly different viruses that can each infect mosquitoes, birds, horses, and humans. Other viruses have host ranges so narrow that they infect only a single species. Measles virus, for instance, can infect only humans. Furthermore, viral infection of multicellular eukaryotes is usually limited to particular tissues. Human cold viruses infect only the cells lining the upper respiratory tract, and the HIV seen in Figure 26.1 binds to receptors present only on certain types of immune cells.

General Features of Viral Replicative Cycles

A viral infection begins when a virus binds to a host cell and the viral genome makes its way inside **(Figure 26.4)**. The mechanism of genome entry depends on the type of virus and the type of host cell. For example, T-even phages use their elaborate tail apparatus to inject DNA into a bacterium (see Figure 26.3d). Other viruses are taken up by endocytosis or, in the case of enveloped viruses, by fusion of the viral envelope with the host's plasma membrane. Once the viral genome is inside, the proteins it encodes can commandeer the host, reprogramming the cell to copy the viral genome and manufacture viral proteins. The host provides the nucleotides for making viral nucleic acids, as well as enzymes, ribosomes, tRNAs, amino acids, ATP, and other components needed for making the viral proteins. Many DNA viruses use the DNA polymerases of the host cell to synthesize new genomes along the templates provided by the viral DNA. In contrast, to replicate their genomes, RNA viruses use virally encoded RNA polymerases that can use RNA as a template. (Uninfected cells generally make no enzymes for carrying out this process.)

Figure 26.4 A simplified viral replicative cycle. A virus is an intracellular parasite that uses the equipment and small molecules of its host cell to replicate. In this simplest of viral cycles, the parasite is a DNA virus with a capsid consisting of a single type of protein.



MAKE CONNECTIONS ➤ *Label each of the straight gray arrows with one word representing the name of the process that is occurring. Review Figure 17.25.*

Animation: Simplified Viral Replicative Cycle

After the viral nucleic acid molecules and capsomeres are produced, they spontaneously self-assemble into new viruses. In fact, researchers can separate the RNA and capsomeres of TMV and then reassemble complete viruses simply by mixing the components together under the right conditions. The simplest type of viral replicative cycle ends with the exit of hundreds or thousands of viruses from the infected host cell, a process that often damages or destroys the cell. Such cellular damage and death, as well as the body's responses to this destruction, cause many of the symptoms associated with viral infections. The viral progeny that exit a cell have the potential to infect additional cells, spreading the viral infection.

There are many variations on the simplified viral replicative cycle we have just described. We will now take a look at some of these variations in bacterial viruses (phages) and animal viruses; later in the chapter, we will consider plant viruses.

Replicative Cycles of Phages

Phages are the best understood of all viruses, although some of them are also among the most complex. Research on phages led to the discovery that some double-stranded DNA viruses can replicate by two alternative mechanisms: the lytic cycle and the lysogenic cycle.

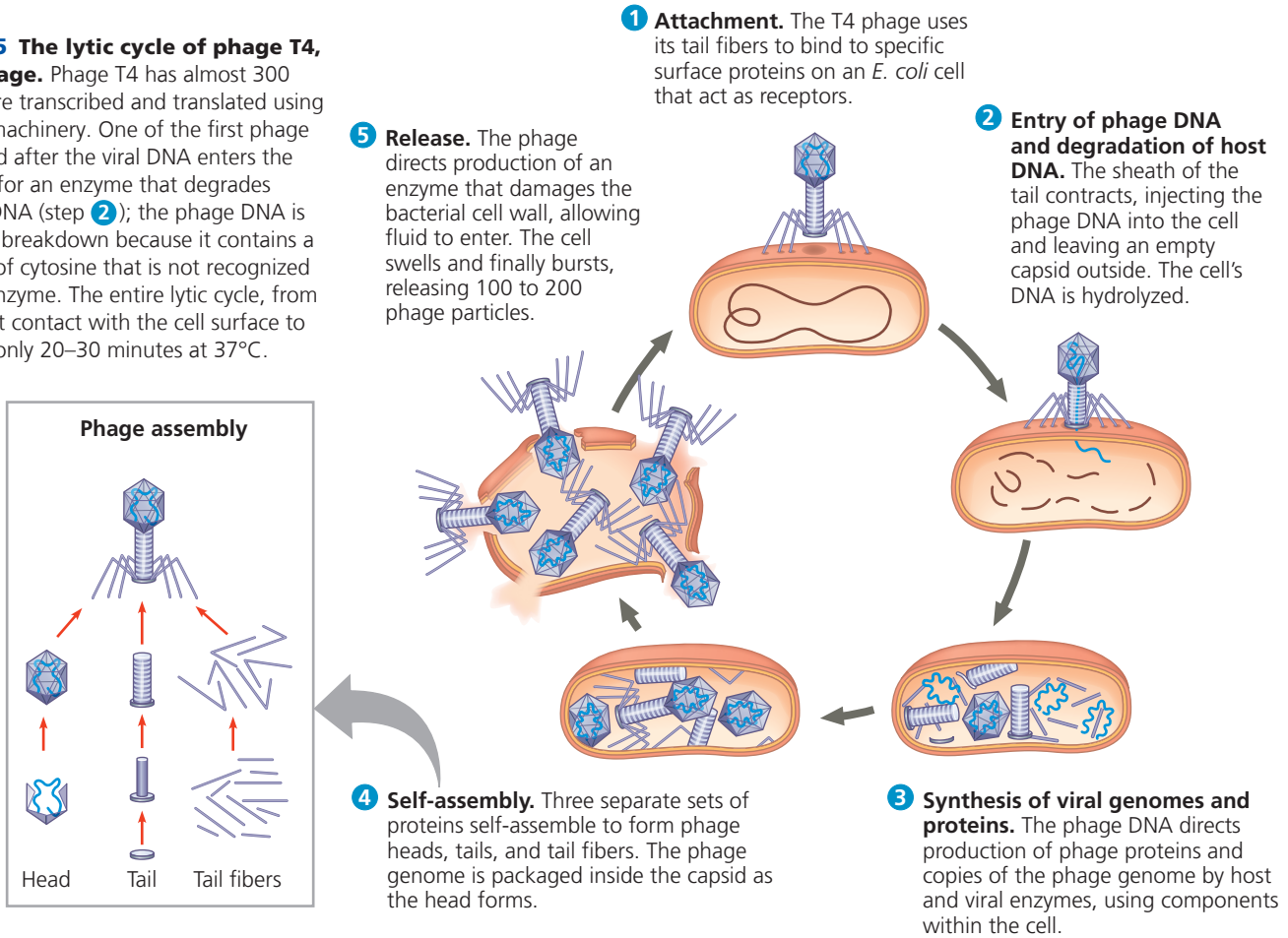
The Lytic Cycle

A phage replicative cycle that culminates in death of the host cell is known as a **lytic cycle**. The term refers to the last stage of infection, during which the bacterium lyses (breaks open) and releases the phages that were produced within the cell. Each of these phages can then infect a healthy cell, and a few successive lytic cycles can destroy an entire bacterial population in just a few hours. A phage that replicates only by a lytic cycle is a **virulent phage**. **Figure 26.5** illustrates the major steps in the lytic cycle of T4, a typical virulent phage.

The Lysogenic Cycle

Instead of lysing their host cells, many phages coexist with them in a state called lysogeny. In contrast to the lytic cycle, which kills the host cell, the **lysogenic cycle** allows replication of the phage genome without destroying the host. Phages capable of using both modes of replicating within a bacterium are called **temperate phages**. A temperate phage called lambda, written with the Greek letter λ, has been widely used in biological research. Phage λ resembles T4, but its tail has only one short tail fiber.

Infection of an *E. coli* cell by phage λ begins when the phage binds to the surface of the cell and injects its linear DNA



➤ Figure 26.5 The lytic cycle of phage T4, a virulent phage. Phage T4 has almost 300 genes, which are transcribed and translated using the host cell's machinery. One of the first phage genes translated after the viral DNA enters the host cell codes for an enzyme that degrades the host cell's DNA (step 2); the phage DNA is protected from breakdown because it contains a modified form of cytosine that is not recognized by the phage enzyme. The entire lytic cycle, from the phage's first contact with the cell surface to cell lysis, takes only 20–30 minutes at 37°C.

MB **Animation: Phage Lytic Cycle**

genome **(Figure 26.6)**. Within the host, the λ DNA molecule forms a circle. What happens next depends on the replicative mode: lytic cycle or lysogenic cycle. During a lytic cycle, the viral genes immediately turn the host cell into a λ-producing factory, and the cell soon lyses and releases its virus progeny. During a lysogenic cycle, however, the λ DNA molecule is incorporated into a specific site on the *E. coli* chromosome by viral proteins that break both circular DNA molecules and join them to each other. When integrated into the bacterial chromosome in this way, the viral DNA is known as a **prophage**. One prophage gene codes for a protein that prevents transcription of most of the other prophage genes. Thus, the phage genome is mostly silent within the bacterium. Every time the *E. coli* cell prepares to divide, it replicates the phage DNA along with its own chromosome such that each daughter cell inherits a prophage. A single infected cell can quickly give rise to a large population of bacteria carrying the virus in prophage form. This mechanism enables viruses to propagate without killing the host cells on which they depend.

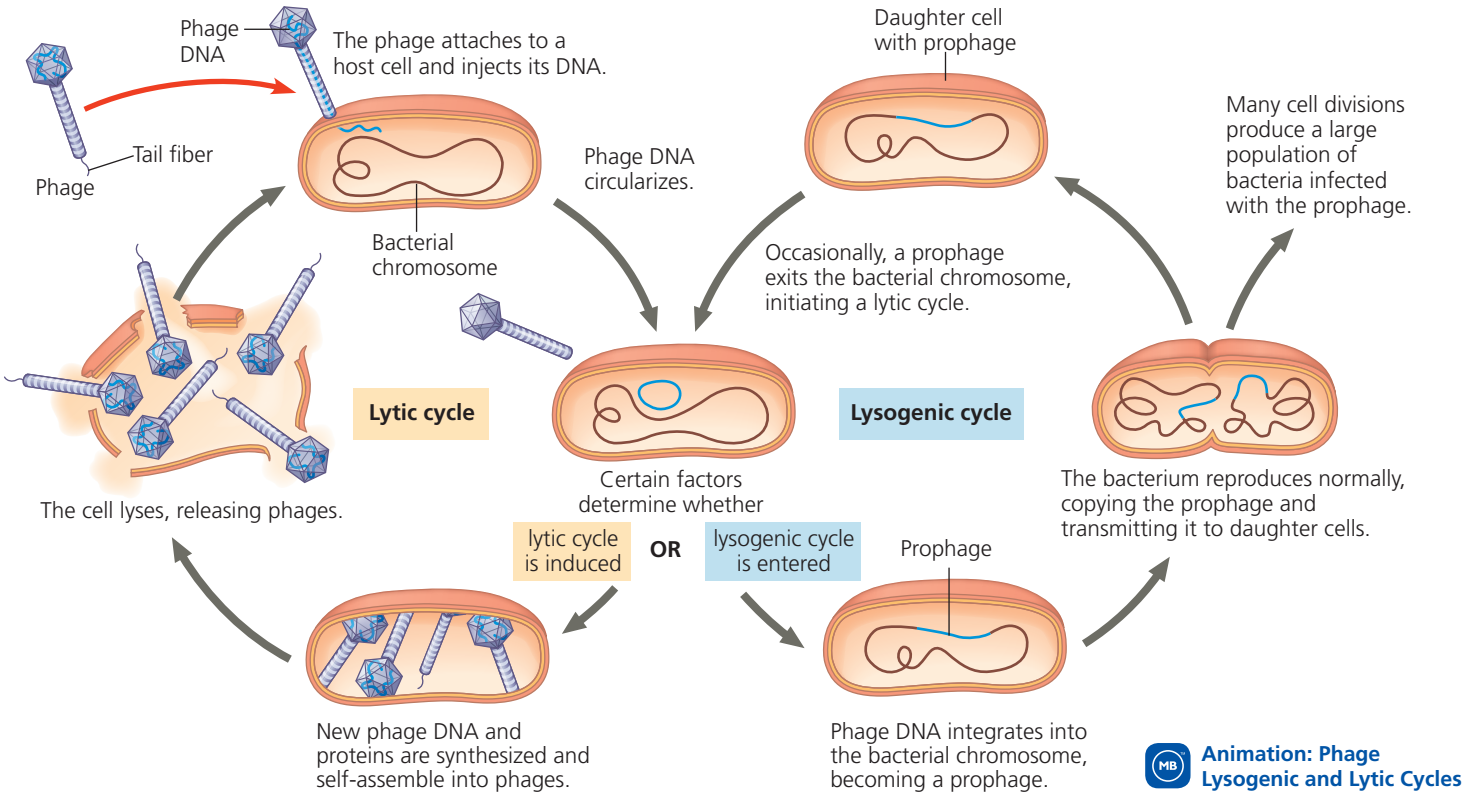
The term *lysogenic* signifies that prophages are capable of generating active phages that lyse their host cells. This occurs when the λ genome (or that of another temperate phage) is induced to exit the bacterial chromosome and initiate a lytic cycle. An environmental signal, such as a certain chemical or high-energy radiation, usually triggers the switchover from the lysogenic to the lytic mode.

In addition to the gene for the viral protein that prevents transcription, a few other prophage genes may be expressed during lysogeny. Expression of these genes may alter the host's phenotype, a phenomenon that can have important medical significance. For example, the three species of bacteria that cause the human diseases diphtheria, botulism, and scarlet fever would not be so harmful to humans without certain prophage genes that cause the host bacteria to make toxins. And the difference between the *E. coli* strain in our intestines and the O157:H7 strain that has caused several deaths by food poisoning appears to be the presence of toxin genes of prophages in the O157:H7 strain.

Bacterial Defenses Against Phages

After reading about the lytic cycle, you may have wondered why phages haven't exterminated all bacteria. Lysogeny is one major reason why bacteria have been spared from extinction caused by phages. Bacteria also have their own defenses against phages. First, natural selection favors bacterial mutants with surface proteins that are no longer recognized as receptors by a particular type of phage. Second, when phage DNA does enter a bacterium, the DNA

Figure 26.6 The lytic and lysogenic cycles of phage λ, a temperate phage. After entering the bacterial cell and circularizing, the λ DNA can immediately initiate the production of a large number of progeny phages (lytic cycle) or integrate into the bacterial chromosome (lysogenic cycle). In most cases, phage λ follows the lytic pathway, which is similar to that detailed in Figure 26.5. However, once a lysogenic cycle begins, the prophage may be carried in the host cell's chromosome for many generations. Phage λ has one main tail fiber, which is short.



Animation: Phage Lysogenic and Lytic Cycles

often is identified as foreign and cut up by cellular enzymes called **restriction enzymes**, which are so named because they *restrict* a phage's ability to replicate within the bacterium. (Restriction enzymes are used in molecular biology and DNA cloning techniques; see Concept 19.1.) The bacterium's own DNA is methylated in a way that prevents attack by its own restriction enzymes. A third defense is a system present in both bacteria and archaea called the *CRISPR-Cas system*.

The CRISPR-Cas system was discovered during a study of repetitive DNA sequences present in the genomes of many prokaryotes. These sequences, which puzzled scientists, were named clustered regularly interspaced short palindromic repeats (CRISPRs) because each sequence read the same forward and backward (a palindrome), with different stretches of "spacer DNA" in between the repeats. At first, scientists assumed the spacer DNA sequences were random and meaningless, but analysis by several research groups showed that each spacer sequence corresponded to DNA from a particular phage that had infected the cell. Further studies revealed that particular nuclease proteins interact with the CRISPR region. These nucleases, called Cas (CRISPR-associated) proteins, can identify and cut phage DNA, thereby defending the bacterium against phage infection.

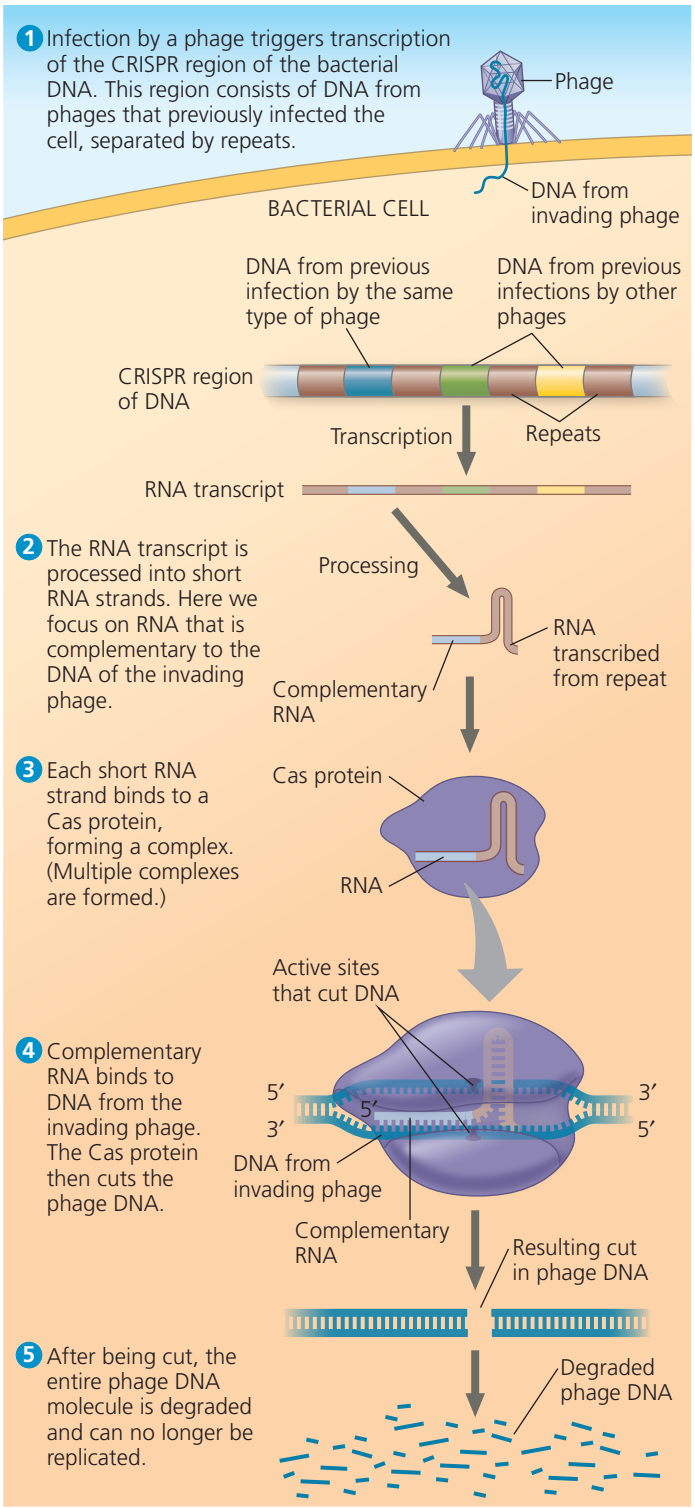
When a phage infects a bacterial cell that has the CRISPR-Cas system, the DNA of the invading phage is integrated into the genome between two repeat sequences. If the cell survives the infection, any further attempt by the same type of phage to infect this cell (or its offspring) triggers transcription of the CRISPR region into RNA molecules **(Figure 26.7)**. These RNAs are cut into pieces and then bound by Cas proteins. The Cas protein uses a portion of the phage-related RNA as a homing device to identify the invading phage DNA and cut it, leading to its destruction. In Concept 19.1, you'll learn how this system is used in the laboratory to alter genes in other cells.

Just as natural selection favors bacteria that have receptors altered by mutation or that have enzymes that cut phage DNA, it also favors phage mutants that can bind to altered receptors or that are resistant to enzymes. Thus, the bacterium-phage relationship is in constant evolutionary flux.

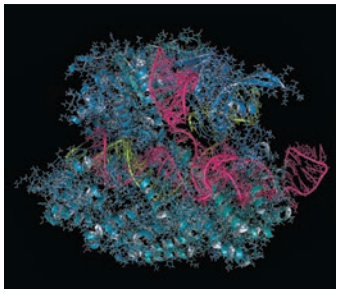
Replicative Cycles of Animal Viruses

Everyone has suffered from viral infections, whether cold sores, influenza, or the common cold. Like all viruses, those that cause illness in humans and other animals can replicate only inside host cells. Many variations on the basic scheme of viral infection and replication are represented among the animal viruses. One key variable is the nature of the viral genome (double- or single-stranded DNA or RNA). Another variable is the presence or absence of a membranous envelope. Whereas few bacteriophages have an envelope or RNA genome, many animal viruses have both. In fact, nearly all

Figure 26.7 The CRISPR-Cas system: a type of bacterial immune system.



Computer model of CRISPR-Cas9 gene editing complex from *Streptococcus pyogenes*



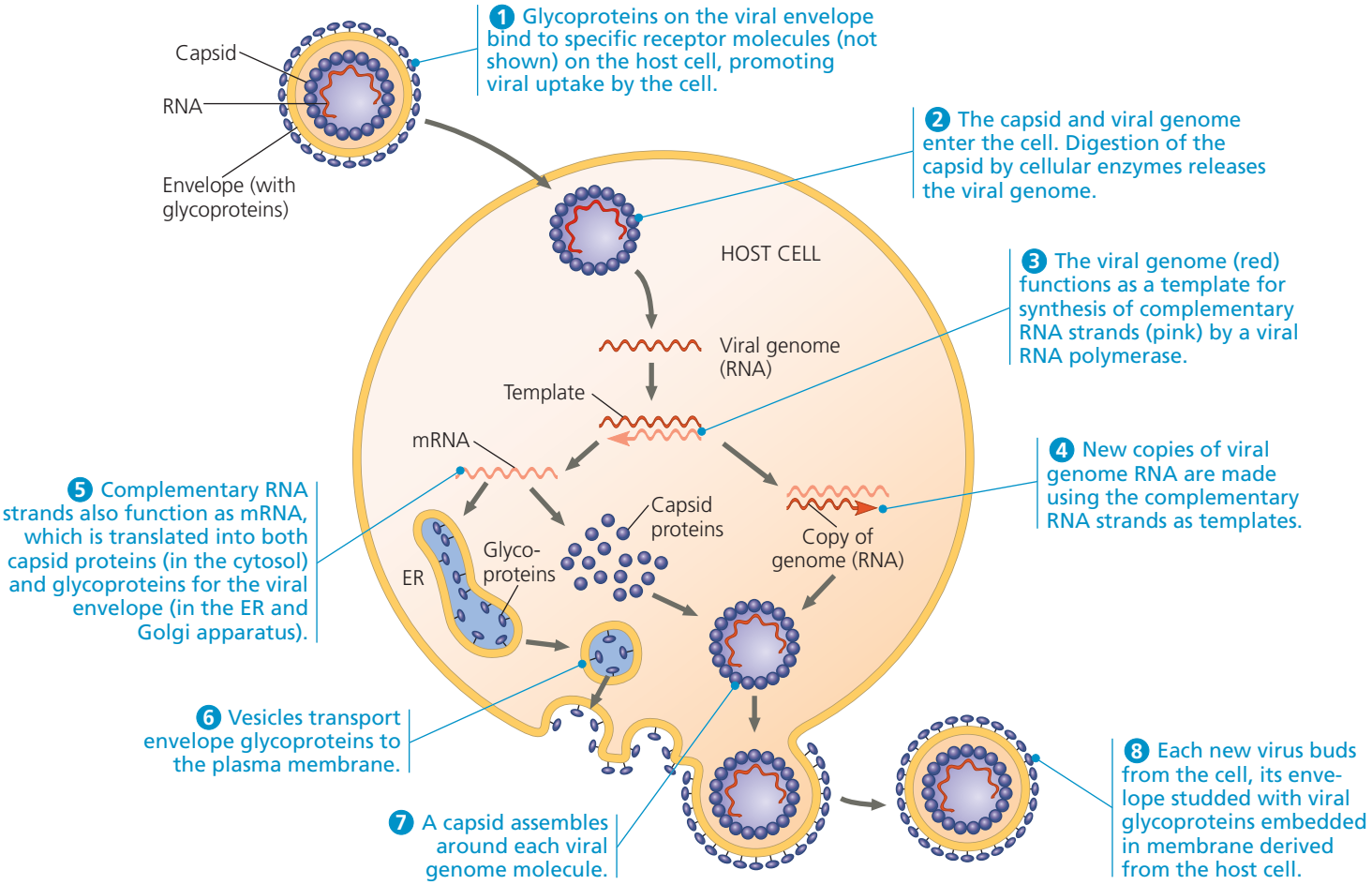
animal viruses with RNA genomes have an envelope, as do some with DNA genomes. Rather than consider all the mechanisms of viral infection and replication, we will focus first on the roles of viral envelopes and then on the functioning of RNA as the genetic material of many animal viruses.

Viral Envelopes

An animal virus equipped with an envelope—that is, a membranous outer layer—uses it to enter the host cell. Protruding from the outer surface of this envelope are viral glycoproteins that bind to specific receptor molecules on the surface of a host cell. **Figure 26.8** outlines the events in the replicative cycle of an enveloped virus with an RNA genome. Ribosomes bound to the endoplasmic reticulum (ER) of the host cell make the protein parts of the envelope glycoproteins; cellular enzymes in the ER and Golgi apparatus then add the sugars. The resulting viral glycoproteins, embedded in membrane derived from the host cell, are transported to the cell surface. In a process much like exocytosis, new viral capsids are wrapped in membrane as they bud from the cell. In other words, the viral envelope is usually derived from the host cell's plasma membrane, although all or most of the molecules of this membrane are specified by viral genes. The enveloped viruses are now free to infect other cells. This replicative cycle does not necessarily kill the host cell, in contrast to the lytic cycles of phages.

Some viruses have envelopes that are not derived from plasma membrane. Herpesviruses, for example, are temporarily cloaked in membrane derived from the nuclear envelope of the host; they then shed this membrane in the cytoplasm and acquire a new envelope made from membrane of the Golgi apparatus. These viruses have a double-stranded DNA genome and replicate within the host cell nucleus, using a combination of viral and cellular enzymes to replicate and transcribe their DNA. In the case of herpesviruses, copies of the viral DNA can remain behind as mini-chromosomes in the nuclei of certain nerve cells. There they remain latent until some sort of physical or emotional stress triggers a new round of active virus production. The infection of other cells by these new viruses causes the blisters characteristic of herpes, such as cold sores or genital sores. Once someone acquires a herpesvirus infection, flare-ups may recur throughout the person's life.

Figure 26.8 The replicative cycle of an enveloped RNA virus. Shown here is a virus with a single-stranded RNA genome that functions as a template for synthesis of mRNA. Some enveloped viruses enter the host cell by fusion of the envelope with the cell's plasma membrane; others enter by endocytosis. For all enveloped RNA viruses, formation of new envelopes for progeny viruses occurs by the mechanism depicted in this figure.



Viral Genetic Material

Table 26.1 shows the common classification system for animal viruses, which is based on their genetic material: double- or single-stranded DNA, or double- or single-stranded RNA. Although some phages and most plant viruses are RNA viruses, the broadest variety of RNA genomes is found among the viruses that infect animals. There are three types of single-stranded RNA genomes found in animal viruses (classes IV–VI in Table 26.1). The genome of class IV viruses can directly serve as mRNA and thus can be translated into viral protein immediately after infection. Figure 26.8 shows a virus of class V, in which the RNA genome serves instead as a *template* for mRNA synthesis. The RNA genome is transcribed into complementary RNA strands, which function both as mRNA and as templates for the synthesis of additional copies of genomic RNA. All viruses that use an RNA genome as a template for mRNA transcription require RNA → RNA synthesis. These viruses use a viral enzyme capable of carrying out this process; there are no such enzymes in most cells. The enzyme used in this process is encoded by the viral genome, and after its synthesis the protein is packaged during viral self-assembly with the genome inside the viral capsid.

Table 26.1 Classes of Animal Viruses

Class/Family	Envelope?	Examples That Cause Human Diseases
I. Double-Stranded DNA (dsDNA)		
Adenovirus (see Figure 26.3b)	No	Respiratory viruses
Papillomavirus	No	Warts, cervical cancer
Polyomavirus	No	Tumors
Herpesvirus	Yes	Herpes simplex I and II (cold sores, genital sores); varicella zoster (shingles, chicken pox); Epstein-Barr virus (mononucleosis, Burkitt's lymphoma)
Poxvirus	Yes	Smallpox virus; cowpox virus
II. Single-Stranded DNA (ssDNA)		
Parvovirus	No	B19 parvovirus (mild rash)
III. Double-Stranded RNA (dsRNA)		
Reovirus	No	Rotavirus (diarrhea); Colorado tick fever virus
IV. Single-Stranded RNA (ssRNA); Serves as mRNA		
Picornavirus	No	Rhinovirus (common cold); poliovirus; hepatitis A virus; other intestinal viruses
Coronavirus	Yes	Severe acute respiratory syndrome (SARS); Middle East respiratory syndrome (MERS)
Flavivirus	Yes	Zika virus (see Figure 26.10c); yellow fever virus; dengue virus; West Nile virus; hepatitis C virus
Togavirus	Yes	Chikungunya virus (see Figure 26.10b); rubella virus; equine encephalitis viruses
V. ssRNA; Serves as Template for mRNA Synthesis		
Filovirus	Yes	Ebola virus (hemorrhagic fever; see Figure 26.10a)
Orthomyxovirus	Yes	Influenza virus (see Figure 26.3c)
Paramyxovirus	Yes	Measles virus; mumps virus
Rhabdovirus	Yes	Rabies virus
VI. ssRNA; Serves as Template for DNA Synthesis		
Retrovirus	Yes	Human immunodeficiency virus (HIV/AIDS; see Figure 26.9); RNA tumor viruses (leukemia)

The RNA animal viruses with the most complicated replicative cycles are the **retroviruses** (class VI). These viruses have an enzyme called **reverse transcriptase** that transcribes an RNA template into DNA, an RNA → DNA information flow that is the opposite of the usual direction. This unusual phenomenon is the source of the name retroviruses (*retro* means "backward"). Of particular medical importance is **HIV (human immunodeficiency virus)**, the retrovirus shown in Figure 26.1 that causes **AIDS (acquired immunodeficiency syndrome)**. HIV and other retroviruses are enveloped viruses that contain two identical molecules of single-stranded RNA and two molecules of reverse transcriptase.

The HIV replicative cycle (traced in **Figure 26.9**) is typical of a retrovirus. After HIV enters a host cell, its reverse transcriptase molecules are released into the cytoplasm, where they catalyze synthesis of viral DNA. The newly made viral DNA then enters the cell's nucleus and integrates into the DNA of a chromosome. The integrated viral DNA, called a **provirus**, never leaves the host's genome, remaining a permanent resident of the cell. (Recall that a prophage, in contrast, leaves the host's genome at the start of a lytic cycle.) The RNA polymerase of the host transcribes the proviral DNA into RNA molecules, which can function both as mRNA for the synthesis of viral proteins and as genomes for the new viruses that will be assembled and released from the cell. In Concept 47.4, we describe how HIV causes the deterioration of the immune system that occurs in AIDS.

Evolution of Viruses

EVOLUTION We began this chapter by asking whether or not viruses are alive. Viruses do not really fit our definition of living organisms. An isolated virus is biologically inert, unable to replicate its genes or regenerate its own ATP. Yet it has a genetic program written in the universal language of life. Do we think of viruses as nature's most complex associations of molecules or as the simplest forms of life? Either way, we must bend our usual definitions. Although viruses cannot replicate or carry out metabolic activities independently, their use of the genetic code makes it hard to deny their evolutionary connection to the living world.

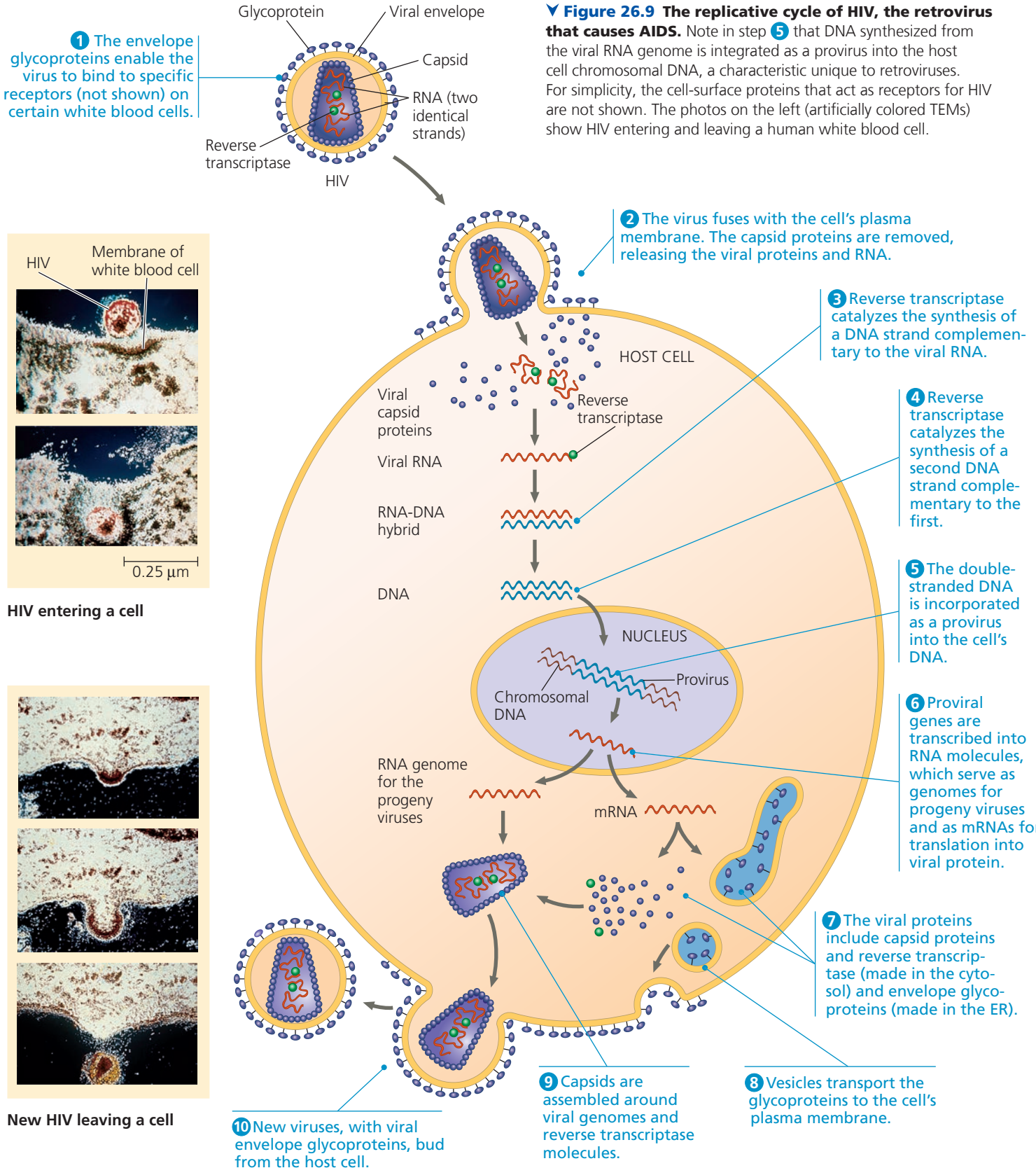


Figure 26.9 The replicative cycle of HIV, the retrovirus that causes AIDS. Note in step 5 that DNA synthesized from the viral RNA genome is integrated as a provirus into the host cell chromosomal DNA, a characteristic unique to retroviruses. For simplicity, the cell-surface proteins that act as receptors for HIV are not shown. The photos on the left (artificially colored TEMs) show HIV entering and leaving a human white blood cell.

MAKE CONNECTIONS ➤ *Describe what is known about binding of HIV to immune system cells. (See Figure 8.8.) How was this discovered?*

Animation: Retrovirus (HIV) Replicative Cycle

How did viruses originate? Viruses have been found that infect every form of life—not just bacteria, animals, and plants, but also archaea, fungi, and algae and other protists. Because they depend on cells for their own propagation, it seems likely that viruses are not the descendants of precellular forms of life but evolved—possibly multiple times—*after* the first cells appeared. Most molecular biologists favor the hypothesis that viruses originated from naked bits of cellular nucleic acids that moved from one cell to another, perhaps via injured cell surfaces. The evolution of genes coding for capsid proteins may have allowed viruses to bind cell membranes, thus facilitating the infection of uninjured cells.

Candidates for the original sources of viral genomes include plasmids and transposons. *Plasmids* are small, circular DNA molecules found in bacteria and in the unicellular eukaryotes called yeasts. Plasmids exist apart from and can replicate independently of the bacterial chromosome and are occasionally transferred between cells. *Transposons* are DNA segments that can move from one location to another within a cell's genome. Thus, plasmids, transposons, and viruses all share an important feature: They are *mobile genetic elements*. (We'll discuss plasmids in more detail in Concepts 19.1 and 27.2 and transposons in Concept 20.4.)

Consistent with this notion of pieces of DNA shuttling from cell to cell is the observation that a viral genome can have more in common with the genome of its host than with the genomes of viruses that infect other hosts. Indeed, some viral genes are essentially identical to genes of the host.

The debate about the origin of viruses was reinvigorated about 15 years ago by reports of one of the largest viruses yet discovered: Mimivirus is a double-stranded DNA (dsDNA) virus with an icosahedral capsid that is 400 nm in diameter, the size of a small bacterium. Its genome contains 1.2 million bases (Mb)—about 100 times as many as the influenza virus genome—and an estimated 1,000 genes. Perhaps the most surprising aspect of mimivirus, however, was that its genome included genes previously found only in cellular genomes. Some of these genes code for proteins involved in translation, DNA repair, protein folding, and polysaccharide synthesis. Whether mimivirus evolved *before* the first cells and then developed an exploitative relationship with them or evolved more recently and simply scavenged genes from its hosts is not yet settled. Since 2013 several even larger viruses have been discovered that cannot be classified with any existing known virus. One such virus is 1 μm (1,000 nm) in diameter, with a dsDNA genome of around 2–2.5 Mb, larger than that of some small eukaryotes. What's more, over 90% of its 2,000 or so genes are unrelated to cellular genes, inspiring the name it was given, pandoravirus. A second virus, called *Pithovirus sibericum*, with a diameter of 1.5 μm and 500 genes, was discovered in permanently frozen soil in Siberia. This virus, once thawed, was able to infect an amoeba after being frozen for 30,000 years! How these and all other viruses fit in the tree of life is an intriguing, unresolved question.

The ongoing evolutionary relationship between viruses and the genomes of their host cells is an association that continues to make viruses very useful experimental systems in molecular biology. Knowledge about viruses also allows many practical applications, since viruses have a tremendous impact on all organisms through their ability to cause disease.

CONCEPT CHECK 26.2

1. Compare the effect on the host cell of a lytic (virulent) phage and a lysogenic (temperate) phage.
2. **MAKE CONNECTIONS ➤** Compare the CRISPR-Cas system to the miRNA system discussed in Concept 18.3, including their mechanisms and their functions.
3. **MAKE CONNECTIONS ➤** The RNA virus in Figure 26.8 has a viral RNA polymerase that functions in step 3 of the virus's replicative cycle. Compare this with a cellular RNA polymerase in terms of template and overall function (see Figure 17.10).
4. Why is HIV called a retrovirus?
5. **VISUAL SKILLS ➤** Looking at Figure 26.9, imagine you are a researcher trying to combat HIV infection. What molecular processes could you attempt to block?

For suggested answers, see Appendix A.

CONCEPT 26.3

Viruses and prions are formidable pathogens in animals and plants

Diseases caused by viral infections afflict humans, agricultural crops, and livestock worldwide. Other smaller, less complex entities known as prions also cause disease in animals. We'll first discuss animal viruses.

Viral Diseases in Animals

A viral infection can produce symptoms by a number of different routes. Viruses may damage or kill cells by causing the release of hydrolytic enzymes from lysosomes. Some viruses cause infected cells to produce toxins that lead to disease symptoms, and some have molecular components that are toxic, such as envelope proteins. How much damage a virus causes depends partly on the ability of the infected tissue to regenerate by cell division. People usually recover completely from colds because the epithelium of the respiratory tract, which the viruses infect, can efficiently repair itself. In contrast, damage inflicted by poliovirus to mature nerve cells is permanent because these cells do not divide and usually cannot be replaced. Many of the temporary symptoms associated with viral infections, such as fever and body aches, actually result from the body's own efforts to defend itself against infection rather than from cell death caused by the virus.

The immune system is a critical part of the body's natural defenses (see Chapter 47). It is also the basis for the major medical tool used to prevent viral infections—vaccines. A **vaccine** is a harmless derivative of a pathogen that stimulates

the immune system to mount defenses against the harmful pathogen. Smallpox, a viral disease that was once a devastating scourge in many parts of the world, was eradicated by a vaccination program carried out by the World Health Organization (WHO). The very narrow host range of the smallpox virus—it infects only humans—was a critical factor in the success of this program. Similar worldwide vaccination campaigns are currently under way to eradicate polio and measles. Effective vaccines are also available to protect against rubella, mumps, hepatitis B, and a number of other viral diseases.

Although vaccines can prevent some viral illnesses, medical care can do little, at present, to cure most viral infections once they occur. The antibiotics that help us recover from bacterial infections are powerless against viruses. Antibiotics kill bacteria by inhibiting enzymes specific to bacteria but have no effect on eukaryotic or virally encoded enzymes. However, the few enzymes that are encoded only by viruses have provided targets for other drugs. Most antiviral drugs resemble nucleosides and thus interfere with viral nucleic acid synthesis. One such drug is acyclovir, which impedes herpesvirus replication by inhibiting the viral polymerase that synthesizes viral DNA but not the eukaryotic one. Similarly, azidothymidine (AZT) curbs HIV replication by interfering with the synthesis of DNA by reverse transcriptase. In the past 20 years, much effort has gone into developing drugs to treat HIV. Currently, multidrug treatments, sometimes called "cocktails," are considered to be most effective. Such treatments commonly include a combination of two nucleoside mimics and a protease inhibitor, which interferes with an enzyme required for assembly of the viruses. Another effective treatment involves a drug called maraviroc, which blocks a protein on the surface of human immune cells that helps bind the HIV virus (see Figure 8.8). This drug has also been used successfully to prevent infection in individuals who either have been exposed to, or are at risk of exposure to, HIV.

Emerging Viruses

Viruses that suddenly become apparent are often referred to as *emerging viruses*. HIV, the AIDS virus, is a classic example: This virus appeared in San Francisco in the early 1980s, seemingly out of nowhere, although later studies uncovered a case in the Belgian Congo in 1959. A number of other dangerous emerging viruses cause encephalitis, inflammation of the brain. One example is the West Nile virus, which appeared in North America in 1999 and has spread to all 48 contiguous states in the United States, by now resulting in over 40,000 cases and almost 2,000 deaths.

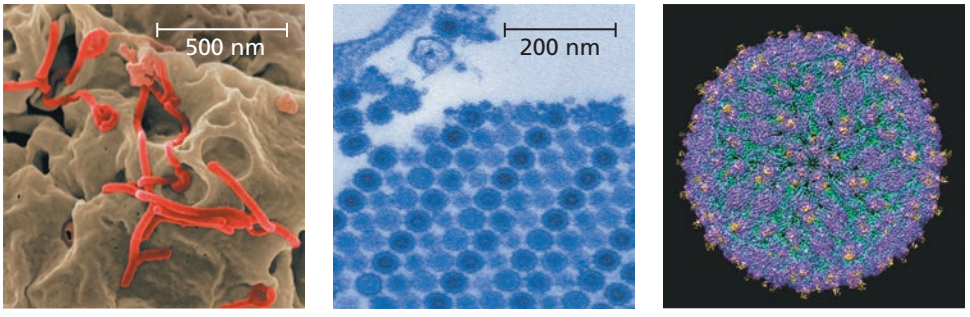
HHMI Video: Interview with Katie Walter

The deadly Ebola virus **(Figure 26.10a)**, recognized initially in 1976 in central Africa, is one of several emerging viruses that cause *hemorrhagic fever*, an often fatal illness characterized by fever, vomiting, massive bleeding, and circulatory system collapse. In 2014, a widespread outbreak of Ebola virus in western Africa caused the World Health Organization to declare an international health emergency. By mid-2015 the outbreak, centered in Guinea, Sierra Leone, and Liberia, had caused over 27,000 illnesses and 11,000 deaths.

Another example is the mosquito-borne virus called chikungunya **(Figure 26.10b)**, which causes an acute illness with fever, rashes, and persistent joint pain. Chikungunya has long been considered a tropical virus, but it has now appeared in northern Italy and southeastern France. A more recently emerging virus is the Zika virus **(Figure 26.10c)**, which caused an outbreak of disease in spring 2015 in Brazil. Although symptoms of Zika are often mild, the outbreak was noticed because infection of pregnant women was correlated with a striking increase in the number of babies born with abnormally small brains, a condition called microcephaly. Zika is a mosquito-borne flavivirus (like West Nile virus) that infects neural cells, posing a particular danger to fetal brain development. Because of the neurological defects associated with Zika and its spread to 28 other countries by early 2016, the World Health Organization declared Zika an international health emergency.

Types of influenza often emerge as outbreaks of illness. In 2009, a widespread outbreak, or **epidemic**, of a flu-like illness appeared in Mexico and the United States. The infectious agent was quickly identified as an influenza virus related to viruses that cause the seasonal flu. This particular virus was named H1N1 for reasons that will be explained shortly. The illness spread rapidly, prompting WHO to declare a global epidemic, or **pandemic**, shortly thereafter. Half a year later, the disease had reached 207 countries, infecting over 600,000 people and killing almost 8,000.

Figure 26.10 Emerging viruses.



(a) Ebola viruses budding from a monkey cell (colorized SEM).

(b) Chikungunya viruses emerging from a cell in the upper left and packing together (colorized TEM).

(c) Computer-generated image of a **Zika virus**, based on a technique called cryo-electron microscopy.

How do such viruses burst on the human scene, giving rise to harmful diseases that were previously rare or even unknown? Three processes contribute to the emergence of viral diseases. The first, and perhaps most important, is the mutation of existing viruses. RNA viruses tend to have an unusually high rate of mutation because viral RNA polymerases do not proofread and correct errors in replicating their RNA genomes. Some mutations change existing viruses into new genetic varieties (strains) that can cause disease, even in individuals who are immune to the ancestral virus. For instance, seasonal flu epidemics are caused by new strains of influenza virus genetically different enough from earlier strains that people have little immunity to them. You'll see an example of this process in the **Scientific Skills Exercise**, where you'll analyze genetic changes in variants of the H1N1 flu virus and correlate them with spread of the disease.

A second process that can lead to the emergence of viral diseases is the dissemination of a viral disease from a small, isolated human population. For instance, AIDS went unnamed and virtually unnoticed for decades before it began to spread around the world. In this case, technological and social factors, including affordable international travel, blood transfusions, sexual promiscuity, and the abuse of intravenous drugs, allowed a previously rare human disease to become a global scourge.

Interview with David Satcher: The role of the CDC in recognizing AIDS and in public health

A third source of new viral diseases in humans is the spread of existing viruses from other animals. Scientists estimate that about three-quarters of new human diseases originate in this way. Animals that harbor and can transmit a particular virus but are generally unaffected by it are said to act as a natural reservoir for that virus. For example, the H1N1 virus that caused the 2009 flu pandemic mentioned earlier was likely passed to humans from pigs; for this reason, the disease it caused was originally called "swine flu."

In general, flu epidemics provide an instructive example of the effects of viruses moving between species. There are three types of influenza virus: types B and C, which infect only humans and have never caused an epidemic, and type A, which infects a wide range of animals, including birds, pigs, horses, and humans. Influenza A strains have caused four major flu epidemics among humans in the last 100 years. The worst was the first one, the "Spanish flu" pandemic of 1918–1919, which killed 40–50 million people, including many World War I soldiers.

Different strains of influenza A are given standardized names; for example, both the strain that caused the 1918 flu and the one that caused the 2009 pandemic flu are called H1N1. The name identifies which forms of two viral surface proteins are present: hemagglutinin (HA) and neuraminidase (NA). There are 16 different types of hemagglutinin, a protein that helps the flu virus attach to host cells, and 9 types of neuraminidase, an enzyme that helps release new virus particles from infected cells. Waterbirds have been found that carry viruses with all possible combinations of HA and NA. Variations of the hemagglutinin protein are used each year to generate vaccines against the strains predicted most likely to occur the next year.

A likely scenario for the 1918 pandemic and others is that the virus mutated as it passed from one host species to another. When an animal like a pig or a bird is infected with more than one strain of flu virus, the different strains can undergo genetic recombination if the RNA molecules making up their genomes mix and match during viral assembly. Pigs were probably the main hosts for recombination that led to the 2009 flu virus, which turns out to contain sequences from bird, pig, and human flu viruses. Coupled with mutation, these reassortments can lead to the emergence of a viral strain capable of infecting human cells. People who have never been exposed to that particular strain before will lack immunity, and the recombinant virus has the potential to be highly pathogenic. If such a flu virus recombines with viruses that circulate widely among humans, it may acquire the ability to spread easily from person to person, dramatically increasing the potential for a major human outbreak.

The many avian flu viruses carried by wild and domestic birds pose a potential long-term threat. A case in point is an H5N1 virus; the first transmission of H5N1 from birds to humans was documented in Hong Kong in 1997. Since then, the overall mortality rate due to H5N1 has been greater than 50% of those infected—an alarming number. Also, the host range of H5N1 is expanding, which provides increasing chances for reassortment between different strains. If the H5N1 avian flu virus evolves so that it can spread easily from person to person, it could represent a major global health threat akin to that of the 1918 pandemic.

How easily could this happen? In 2011, scientists working with ferrets, small mammals that are animal models for human flu, found out that only a few mutations of the avian flu virus would allow infection of cells in the human nasal cavity and windpipe. Furthermore, when the scientists transferred nasal swabs serially from ferret to ferret, the virus became transmissible through the air. Reports of this startling discovery at a scientific conference ignited a firestorm of debate about whether to publish the results and led to an ongoing reevaluation of the federal policies governing this type of experiment in the United States. The risks of doing this type of research (what if the new virus escapes or the procedure falls into the hands of bioterrorists?) must be considered in relation to the risks of not doing it—the possibility that we will be unable to combat new, more transmissible viruses because we lack an understanding of how they develop.

As we have seen, emerging viruses are generally not new; rather, they are existing viruses that mutate, disseminate more widely in the current host species, or spread to new host species. Changes in host behavior or environmental changes

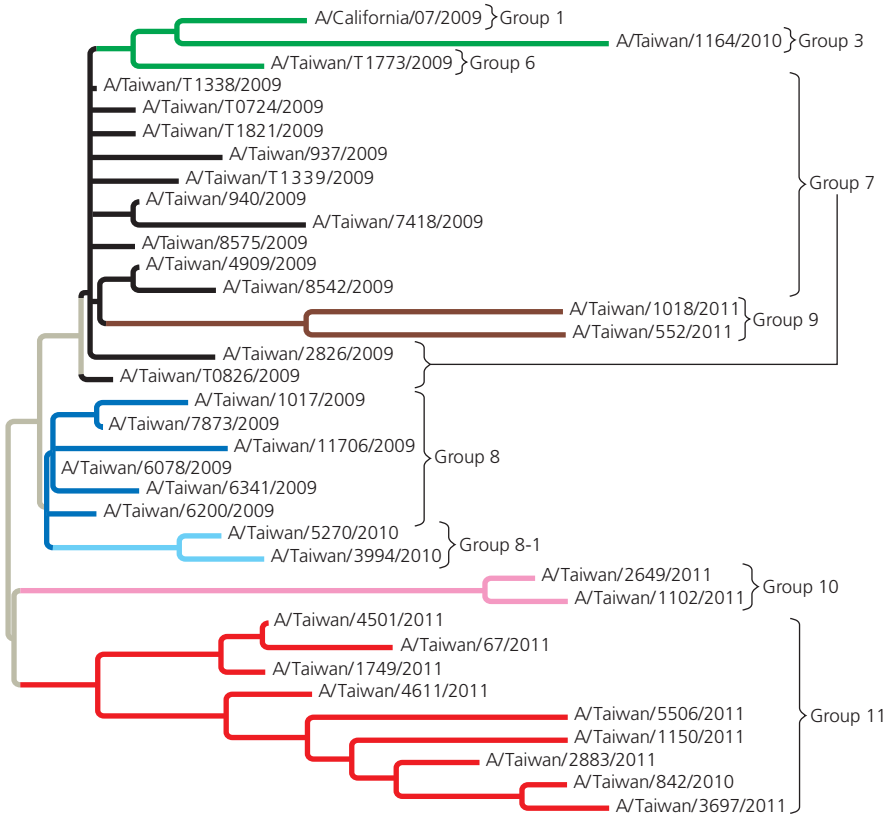
SCIENTIFIC SKILLS EXERCISE

Analyzing a Sequence-Based Phylogenetic Tree to Understand Viral Evolution

How Can Sequence Data Be Used to Track Flu Virus Evolution? In 2009, an influenza A H1N1 virus caused a pandemic, and the virus has continued to resurface in outbreaks across the world. Researchers in Taiwan were curious about why the virus kept appearing despite widespread flu vaccine initiatives. They hypothesized that newly evolved variant strains of the H1N1 virus were able to evade human immune system defenses. To test this hypothesis, they needed to determine if each wave of flu infection was caused by a different H1N1 variant strain.

H1N1 flu vaccination.

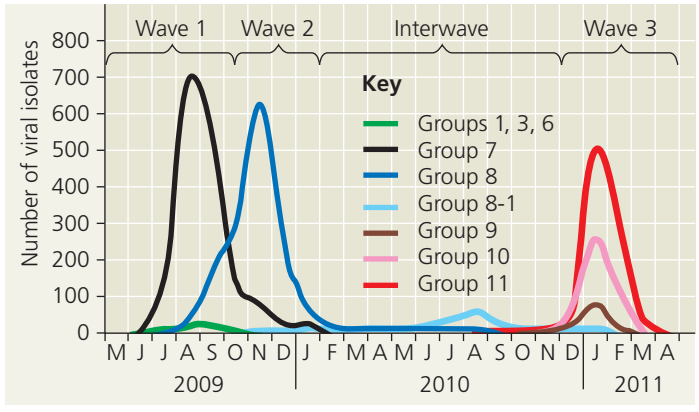
How the Experiment Was Done Scientists obtained the genome sequences for 4,703 virus isolates collected from patients with H1N1 flu in Taiwan. They compared the sequences in different strains for the viral hemagglutinin (HA) gene, and based on mutations that had occurred, arranged the isolates into a phylogenetic tree (see Figure 22.5 for information on how to read phylogenetic trees).



Data from the Experiment In the phylogenetic tree, each branch tip is one variant strain of the H1N1 virus with a unique HA gene sequence. The tree is a way to visualize a working hypothesis about the evolutionary relationships between H1N1 variants.

INTERPRET THE DATA

1. The phylogenetic tree shows the hypothesized evolutionary relationship between the variant strains of H1N1 virus. The more closely connected two variants are, the more alike they are in terms of HA gene sequence. Each fork in a branch, called a node, shows where two lineages separate due to different accumulated mutations. The length of the branches is a measure of how many sequence differences there are between the variants, indicating how distantly related they are. Referring to the phylogenetic tree, which variants are more closely related to each other: A/Taiwan/1018/2011 and A/Taiwan/552/2011 or A/Taiwan/1018/2011 and A/Taiwan/8542/2009? Explain your answer.
2. The scientists arranged the branches into groups made up of one ancestral variant and all of its descendant, mutated variants. They are color-coded in the figure. Using group 11 as an example, trace the lineage of its variants. (a) Do all of the nodes have the same number of branches or branch tips? (b) Are all of the branches in the group the same length? (c) What do these results indicate?
3. The graph at the lower left shows the number of isolates collected (each from an ill patient) on the *y*-axis and the month and year that the isolates were collected on the *x*-axis. Each group of variants is plotted separately with a line color that matches the tree diagram. (a) Which group of variants was the earliest to cause the first wave of H1N1 flu in over 100 patients in Taiwan? (b) After a group of variants had a peak number of infections, did members of that same group cause another (later) wave of infection? (c) One variant in group 1 (green, uppermost branch) was used to make a vaccine that was distributed very early in the pandemic. Based on the graphed data, does it look like the vaccine was effective?
4. Groups 9, 10, and 11 all had H1N1 variants that caused a large number of infections at the same time in Taiwan. Does this mean that the scientists' hypothesis, that new variants cause new waves of infection, was incorrect? Explain your answer.



Scientists graphed the number of isolates by the month and year of isolate collection to show the period in which each viral variant was actively causing illness in people.

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Data from J.-R. Yang et al., New variants and age shift to high fatality groups contribute to severe successive waves in the 2009 influenza pandemic in Taiwan, *PLoS ONE* 6(11): e28288 (2011).

can increase the viral traffic responsible for emerging diseases. For instance, new roads built through remote areas can allow viruses to spread between previously isolated human populations. Also, the destruction of forests to expand cropland can bring humans into contact with other animals that may host viruses capable of infecting humans. Finally, genetic mutations and changes in host ranges can allow viruses to jump from one species to another. Many viruses, including chikungunya, mentioned earlier, can be transmitted by mosquitoes. A dramatic expansion of the disease caused by chikungunya occurred in the mid-2000s when a mutation in the virus allowed it to infect not only the mosquito species *Aedes aegypti*, but also the related *Aedes albopictus*. Promotion of the use of insecticides and mosquito netting over beds are crucial tools in public health attempts to prevent diseases carried by mosquitoes **(Figure 26.11)**.

Figure 26.11 Mosquitoes as vectors for disease.
Mosquitoes transmit viruses when they feed on infected blood from one person and then bite other people. Mosquito netting is an important means of preventing infection in affected areas.

Recently, scientists have become concerned about the possible effects of climate change on worldwide viral transmission. Dengue fever, also mosquito-borne, has appeared in Florida and Portugal, regions where it had not been seen before. The possibility that global climate change has allowed mosquito species carrying these viruses to expand their ranges and interact more is troubling because of the increased chance of a mutation allowing a virus species to jump to a new host. This is an area of active research by scientists applying climate change models to what is known about the habitat requirements of mosquito species.

Viral Diseases in Plants

More than 2,000 types of viral diseases of plants are known, and together they account for an estimated annual loss of $15 billion worldwide due to their destruction of agricultural and horticultural crops. Common signs of viral infection include bleached or brown spots on leaves and fruits **(Figure 26.12)**, stunted growth, and damaged flowers or roots, all of which can diminish the yield and quality of crops.

Figure 26.12 Immature tomato infected by a virus.

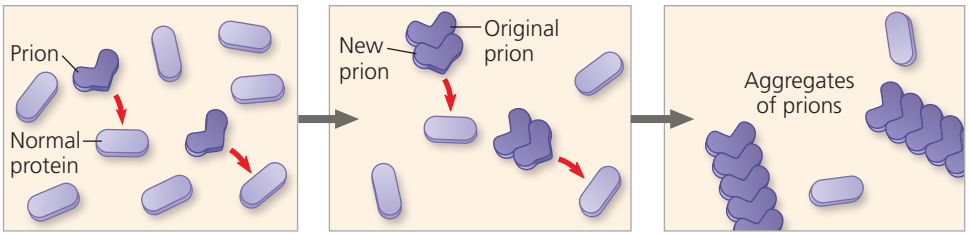
Plant viruses have the same basic structure and mode of replication as animal viruses. Most plant viruses discovered thus far, including tobacco mosaic virus (TMV), have an RNA genome. Many have a helical capsid, like TMV, while others have an icosahedral capsid (see Figure 26.3b).

Viral diseases of plants spread by two major routes. In the first route, called *horizontal transmission*, a plant is infected from an external source of the virus. Because the invading virus must get past the plant's outer protective layer of cells (the epidermis), a plant becomes more susceptible to viral infections if it has been damaged by wind, injury, or herbivores. Herbivores, especially insects, pose a double threat because they can also act as carriers of viruses, transmitting disease from plant to plant. Moreover, farmers and gardeners may transmit plant viruses inadvertently on pruning shears and other tools. The other route of viral infection is *vertical transmission*, in which a plant inherits a viral infection from a parent. Vertical transmission can occur in asexual propagation (for example, through cuttings) or in sexual reproduction via infected seeds.

Once a virus enters a plant cell and begins replicating, viral genomes and associated proteins can spread throughout the plant by means of plasmodesmata, the cytoplasmic connections that penetrate the walls between adjacent plant cells (see Figure 36.19). The passage of viral macromolecules from cell to cell is facilitated by virally encoded proteins that cause enlargement of plasmodesmata. Scientists have not yet devised cures for most viral plant diseases. Consequently, research efforts are focused largely on reducing the transmission of such diseases and on breeding resistant varieties of crop plants.

Prions: Proteins as Infectious Agents

The viruses we have discussed in this chapter are infectious agents that spread diseases, and their genetic material is composed of nucleic acids, whose ability to be replicated is well known. Surprisingly, there are also *proteins* that are known to be infectious. Proteins called **prions** appear to cause a number of degenerative brain diseases in various animal species. These diseases include scrapie in sheep; mad cow disease, which has plagued the European beef industry in recent years; and Creutzfeldt-Jakob disease in humans, which has caused the death of some 175 people in the United Kingdom since 1996. Prions can be transmitted in food, as may occur when people eat prion-laden beef from cattle with mad cow disease. Kuru, another human disease caused by prions, was identified in the early 1900s among the South Fore natives of New Guinea. A kuru epidemic peaked there in the 1960s, puzzling scientists, who at first thought the disease had a genetic basis.



➤ Figure 26.13 Model for how prions propagate. Prions are misfolded versions of normal brain proteins. When a prion contacts a normally folded version of the same protein, it may induce the normal protein to assume the abnormal shape. The resulting chain reaction may continue until high levels of prion aggregation cause cellular malfunction and eventual degeneration of the brain.

Animation: Prions: Characteristics
Animation: Prions: Diseases

Eventually, however, anthropological investigations ferreted out how the disease was spread: ritual cannibalism, a widespread practice among South Fore natives at that time.

Two characteristics of prions are especially alarming. First, prions act very slowly, with an incubation period of at least ten years before symptoms develop. The lengthy incubation period prevents sources of infection from being identified until long after the first cases appear, allowing many more infections to occur. Second, prions are virtually indestructible; they are not destroyed or deactivated by heating to normal cooking temperatures. To date, there is no known cure for prion diseases, and the only hope for developing effective treatments lies in understanding the process of infection.

How can a protein, which cannot replicate itself, be a transmissible pathogen? According to the leading model, a prion is a misfolded form of a protein normally present in brain cells. When the prion gets into a cell containing the normal form of the protein, the prion somehow converts normal protein molecules to the misfolded prion versions. Several prions then aggregate into a complex that can convert other normal proteins to prions, which join the chain **(Figure 26.13)**. Prion aggregation interferes with normal cellular functions and causes disease symptoms. This model was greeted with much skepticism when it was first proposed by Stanley Prusiner in the early 1980s, but it is now widely accepted. Prusiner was awarded the Nobel Prize in 1997 for his work on prions. He has recently proposed that prions are also involved in neurodegenerative diseases such as Alzheimer's and Parkinson's disease. There are many outstanding questions about these small infectious agents.

CONCEPT CHECK 26.3

1. Describe two ways in which a preexisting virus can become an emerging virus.
2. Prions do not contain genetic material and, therefore, cannot replicate themselves. How do they then cause diseases?
3. **WHAT IF? ➤** TMV has been isolated from virtually all commercial tobacco products. Why, then, is TMV infection not an additional hazard for smokers?

For suggested answers, see Appendix A.

26 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 26.1

A virus consists of a nucleic acid surrounded by a protein coat *(pp. 609–611)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

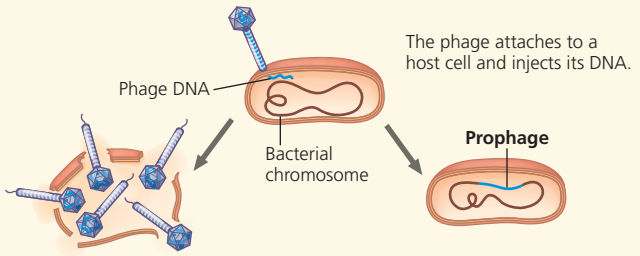
- Researchers discovered viruses in the late 1800s by studying a plant disease, tobacco mosaic disease.
- A **virus** is a small nucleic acid genome enclosed in a protein **capsid** and sometimes a membranous **viral envelope**. The genome may be single- or double-stranded DNA or RNA.

? *Are viruses generally considered living or nonliving? Explain.*

CONCEPT 26.2

Viruses replicate only in host cells *(pp. 611–618)*

- Viruses use enzymes, ribosomes, and small molecules of host cells to synthesize progeny viruses during replication.
- Each type of virus has a characteristic **host range**, affected by whether cell-surface proteins are present that viral surface proteins can bind to.
- **Phages** (viruses that infect bacteria) can replicate by two alternative mechanisms: the **lytic cycle** and the **lysogenic cycle**.



Lytic cycle
- **Virulent** or **temperate phage**
- Destruction of host DNA
- Production of new phages
- Lysis of host cell causes release of progeny phages

Lysogenic cycle
- **Temperate phage** only
- Genome integrates into bacterial chromosome as **prophage**, which (1) is replicated and passed on to daughter cells and (2) can be induced to leave the chromosome and initiate a lytic cycle

- Bacteria have various ways of defending themselves against phage infections, including the CRISPR-Cas system.
- Many animal viruses have an envelope. **Retroviruses** (such as **HIV**) use the enzyme **reverse transcriptase** to copy their RNA genome into DNA, which can be integrated into the host genome as a **provirus**.
- Since viruses can replicate only within cells, they probably evolved after the first cells appeared, perhaps as packaged fragments of cellular nucleic acid.

? *Describe enzymes that are not found in most cells but are necessary for the replication of certain types of viruses.*

CONCEPT 26.3

Viruses and prions are formidable pathogens in animals and plants *(pp. 618–623)*

- Symptoms of viral diseases may be caused by direct viral harm to cells or by the body's immune response. **Vaccines** stimulate the immune system to defend the host against specific viruses.
- An **epidemic**, a widespread outbreak of a disease, can become a **pandemic**, a global epidemic.
- Outbreaks of emerging viral diseases in humans are usually not new, but rather are caused by existing viruses that expand their host territory. The H1N1 2009 flu virus was a new combination of pig, human, and avian viral genes that caused a pandemic. The H5N1 avian flu virus has the potential to cause a high-mortality flu pandemic.
- Viruses enter plant cells through damaged cell walls (horizontal transmission) or are inherited from a parent (vertical transmission).
- **Prions** are slow-acting, virtually indestructible infectious proteins that cause brain diseases in mammals.

? *What aspect of an RNA virus makes it more likely than a DNA virus to become an emerging virus?*

TEST YOUR UNDERSTANDING

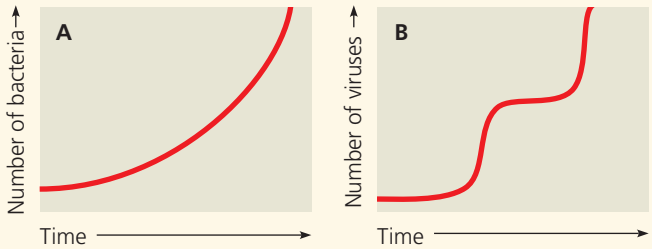
Multiple-choice Self-Quiz questions 1–5 can be found in the Study Area in MasteringBiology.

6. **DRAW IT** Redraw Figure 26.8 to show the replicative cycle of a virus with a single-stranded genome that can function as mRNA (a class IV virus).

PRACTICE TEST
goo.gl/iAsVgL

7. **EVOLUTION CONNECTION** The success of some viruses lies in their ability to evolve rapidly within the host. Such viruses evade the host's defenses by mutating and producing many altered progeny viruses before the body can mount an attack. Thus, the viruses present late in infection differ from those that initially infected the body. Discuss this as an example of evolution in microcosm. Which viral lineages tend to predominate?

8. **SCIENTIFIC INQUIRY** When bacteria infect an animal, the number of bacteria in the body increases in an exponential fashion (graph A). After infection by a virulent animal virus with a lytic replicative cycle, there is no evidence of infection for a while. Then the number of viruses rises suddenly and subsequently increases in a series of steps (graph B). Explain the difference in the curves.



9. **WRITE ABOUT A THEME: ORGANIZATION** While viruses are considered by most scientists to be nonliving, they do show some characteristics of life, including the correlation of structure and function. In a short essay (100–150 words), discuss how the structure of a virus correlates with its function.

10. **SYNTHESIZE YOUR KNOWLEDGE**

Oseltamivir (Tamiflu), an antiviral drug prescribed for influenza, inhibits the enzyme neuraminidase. Explain how this drug could prevent infection in someone exposed to the flu or could shorten the course of flu in an infected patient (the reasons for which it is prescribed).

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

27 Prokaryotes

Figure 27.1 Why is this lake's water pink?

KEY CONCEPTS

Archaea in the genus *Halobacterium*

Masters of Adaptation

At certain times of year, the Laguna Salada de Torrevieja in Spain (the "Salty Lagoon") appears pink **(Figure 27.1)**, a sign that its waters are many times saltier than seawater. Yet despite these harsh conditions, the dramatic color is caused not by minerals or other nonliving sources, but by living things. What organisms can live in such an inhospitable environment, and how do they do it?

The pink color in the Laguna Salada de Torrevieja comes from trillions of prokaryotes in the domains Archaea and Bacteria, including archaea in the genus *Halobacterium*. These archaea have red membrane pigments, some of which capture light energy that is used to drive ATP synthesis. *Halobacterium* species are among the most salt-tolerant organisms on Earth; they thrive in salinities that dehydrate and kill other cells. A *Halobacterium* cell compensates for water lost through osmosis by pumping potassium ions (K^+) into the cell until the ionic concentration inside the cell matches the concentration outside.

Like *Halobacterium*, many other prokaryotes can tolerate extreme conditions. Examples include *Deinococcus radiodurans*, which can survive 3 million rads of radiation (3,000 times the dose fatal to humans), and *Picrophilus oshimae*, which can grow at a pH of 0.03 (acidic enough to dissolve metal). Other prokaryotes live in environments that are too cold or too hot for most other organisms, and some have even been found living in rocks 3.2 km (2 miles) below Earth's surface.

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

Prokaryotic species are also very well adapted to more "normal" habitats—the lands and waters in which most other species are found. Their ability to adapt to a broad range of habitats helps explain why prokaryotes are the most abundant organisms on Earth. Indeed, the number of prokaryotes in a handful of fertile soil is greater than the number of people who have ever lived. In this chapter, we'll examine the adaptations, diversity, and enormous ecological impact of these remarkable organisms.

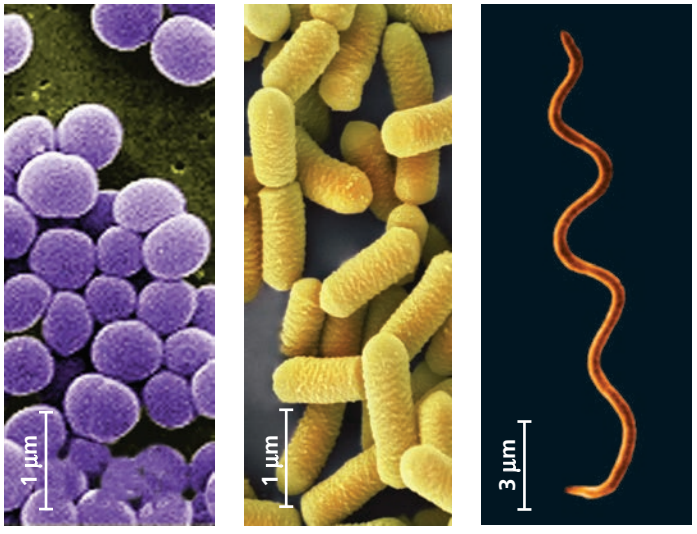
CONCEPT 27.1

Structural and functional adaptations contribute to prokaryotic success

The first organisms to inhabit Earth were prokaryotes that lived 3.5 billion years ago (see Concept 25.3). Throughout their long evolutionary history, prokaryotic populations have been (and continue to be) subjected to natural selection in all kinds of environments, resulting in their enormous diversity today.

We'll begin by describing prokaryotes. Most prokaryotes are unicellular, although the cells of some species remain attached to each other after cell division. Prokaryotic cells typically have diameters of 0.5–5 µm, much smaller than the 10- to 100-µm diameter of many eukaryotic cells. (One notable exception, *Thiomargarita namibiensis*, can be as large as 750 µm in diameter—bigger than a poppy seed.) Prokaryotic cells have a variety of shapes **(Figure 27.2)**.

Figure 27.2 The most common shapes of prokaryotes. **(a)** Cocci (singular, *coccus*) are spherical prokaryotes. They occur singly, in pairs (diplococci), in chains of many cells (streptococci), and in clusters resembling bunches of grapes (staphylococci). **(b)** Bacilli (singular, *bacillus*) are rod-shaped prokaryotes. They are usually solitary, but in some forms the rods are arranged in chains (streptobacilli). **(c)** Spiral prokaryotes include spirilla, which range from comma-like shapes to loose coils, and spirochetes (shown here), which are corkscrew-shaped (colorized SEMs).



(a) Spherical (b) Rod-shaped (c) Spiral

Finally, although they are unicellular and small, prokaryotes are well organized, achieving all of an organism's life functions within a single cell.

Cell-Surface Structures

A key feature of nearly all prokaryotic cells is the cell wall, which maintains cell shape, protects the cell, and prevents it from bursting in a hypotonic environment (see Figure 8.12). In a hypertonic environment, most prokaryotes lose water and shrink away from their wall (plasmolyze). Such water losses can inhibit cell reproduction. Thus, salt can be used to preserve foods because it causes food-spoiling prokaryotes to lose water, preventing them from rapidly multiplying.

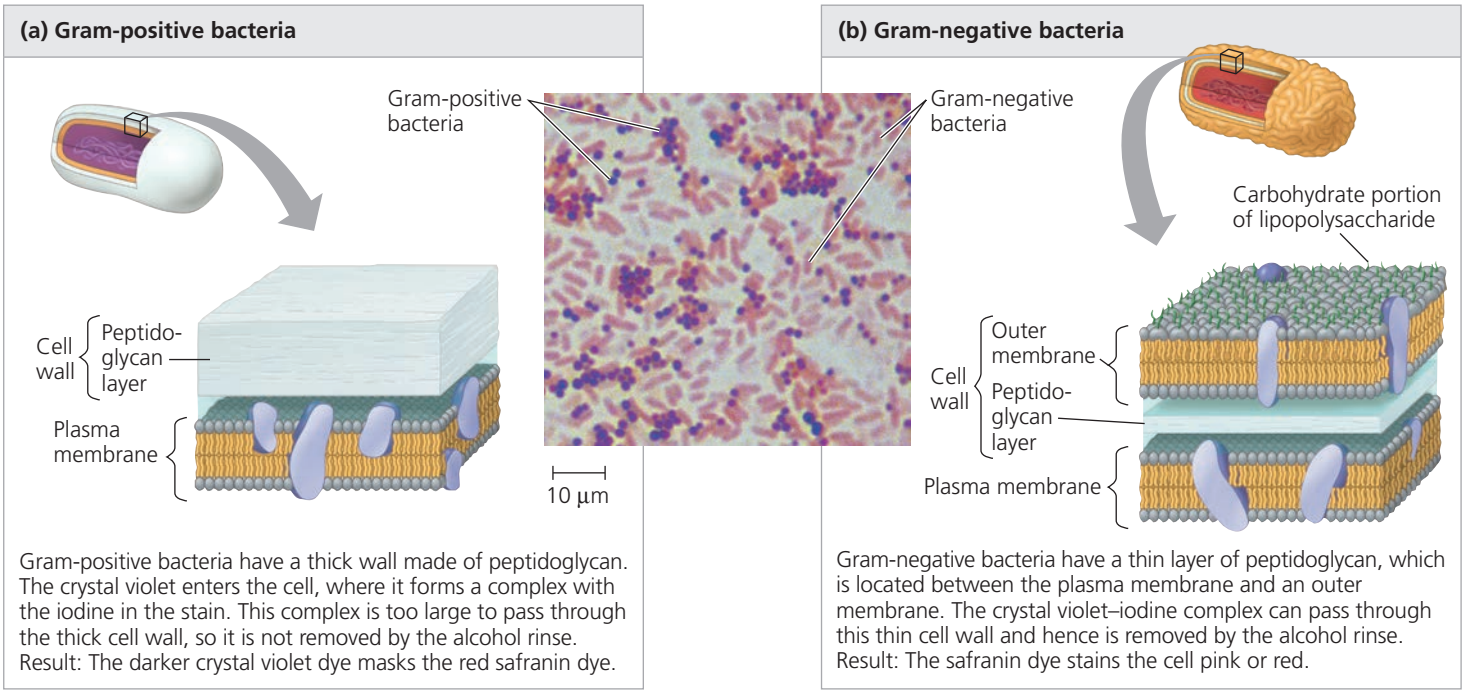
The cell walls of prokaryotes differ in structure from those of eukaryotes. In eukaryotes that have cell walls, such as plants and fungi, the walls are usually made of cellulose or chitin (see Concept 5.2). In contrast, most bacterial cell walls contain **peptidoglycan**, a polymer composed of modified sugars cross-linked by short polypeptides. This molecular fabric encloses the entire bacterium and anchors other molecules that extend from its surface. Archaeal cell walls contain a variety of polysaccharides and proteins but lack peptidoglycan.

Using a technique called the **Gram stain**, developed by the 19th-century Danish physician Hans Christian Gram, scientists can categorize many bacterial species according to differences in cell wall composition. To do this, samples are first stained with crystal violet dye and iodine, then rinsed in alcohol, and finally stained with a red dye such as safranin that enters the cell and binds to its DNA. The structure of a bacterium's cell wall determines the staining response **(Figure 27.3)**. **Gram-positive** bacteria have relatively simple walls composed of a thick layer of peptidoglycan. The walls of **gram-negative** bacteria have less peptidoglycan and are structurally more complex, with an outer membrane that contains lipopolysaccharides (carbohydrates bonded to lipids).

Gram staining is a valuable tool in medicine for quickly determining if a patient's infection is due to gram-negative or to gram-positive bacteria. This information has treatment implications. The lipid portions of the lipopolysaccharides in the walls of many gram-negative bacteria are toxic, causing fever or shock. Furthermore, the outer membrane of a gram-negative bacterium helps protect it from the body's defenses. Gram-negative bacteria also tend to be more resistant than gram-positive species to antibiotics because the outer membrane impedes entry of the drugs. However, certain gram-positive species have virulent strains that are resistant to one or more antibiotics. (Figure 21.14 discusses one example: methicillin-resistant *Staphylococcus aureus*, or MRSA, which can cause lethal skin infections.)

The effectiveness of certain antibiotics, such as penicillin, derives from their inhibition of peptidoglycan cross-linking. The resulting cell wall may not be functional, particularly in

Figure 27.3 Gram staining.



gram-positive bacteria. Such drugs destroy many species of pathogenic bacteria without adversely affecting human cells, which do not have peptidoglycan.

The cell wall of many prokaryotes is surrounded by a sticky layer of polysaccharide or protein. This layer is called a **capsule** if it is dense and well-defined **(Figure 27.4)** or a *slime layer* if it is not as well organized. Both kinds of sticky outer layers enable prokaryotes to adhere to their substrate or to other individuals in a colony. Some capsules and slime layers protect against dehydration, and some shield pathogenic prokaryotes from attacks by their host's immune system.

In another way of withstanding harsh conditions, certain bacteria develop resistant cells called **endospores** when they lack water or essential nutrients **(Figure 27.5)**. The original cell produces a copy of its chromosome and surrounds that copy with a multilayered structure, forming the endospore. Water is removed from the endospore, and its metabolism halts. The original cell then lyses, releasing the endospore. Most endospores are so durable that they can survive in boiling water; killing them requires heating lab equipment to 121°C under high pressure. In less hostile environments, endospores can remain dormant but viable for centuries, able to rehydrate and resume metabolism when their environment improves.

Finally, some prokaryotes stick to their substrate or to one another by means of hairlike appendages called **fimbriae**

Figure 27.4 Capsule. The polysaccharide capsule around this *Streptococcus* bacterium enables the prokaryote to attach to cells in the respiratory tract—in this colorized TEM, a tonsil cell.

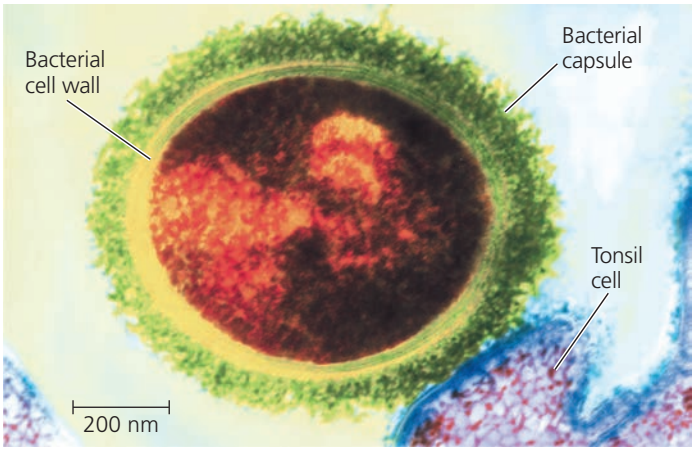


Figure 27.5 An endospore. *Bacillus anthracis*, the bacterium that causes the disease anthrax, produces endospores (TEM). An endospore's protective, multilayered coat helps it survive in the soil for years.

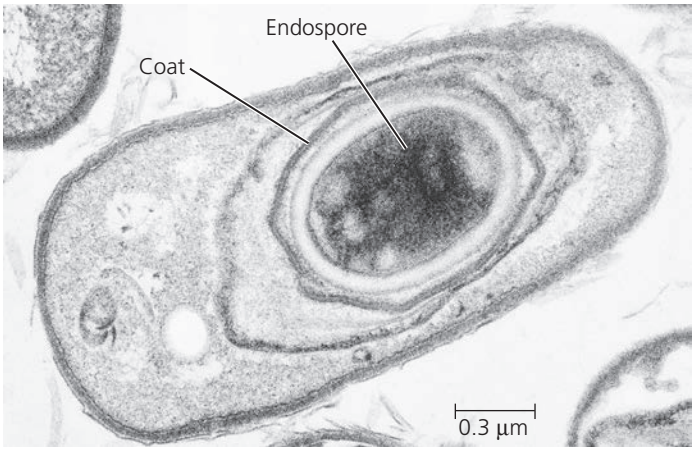
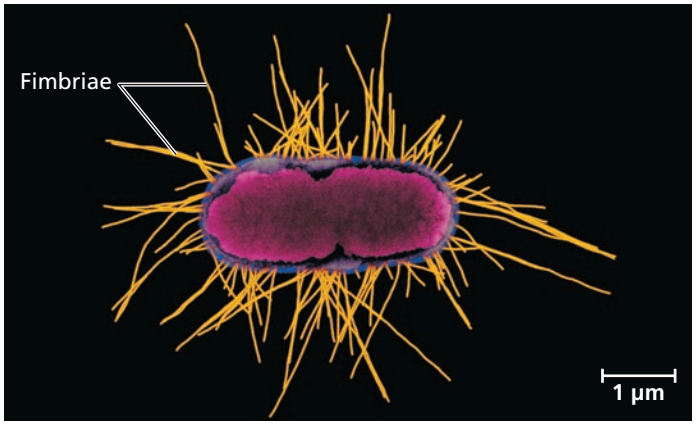


Figure 27.6 Fimbriae. These numerous protein-containing appendages enable some prokaryotes to attach to surfaces or to other cells (colorized TEM).



(singular, *fimbria*) **(Figure 27.6)**. For example, the bacterium that causes gonorrhea, *Neisseria gonorrhoeae*, uses fimbriae to fasten itself to the mucous membranes of its host. Fimbriae are usually shorter and more numerous than **pili** (singular, *pilus*), appendages that pull two cells together prior to DNA transfer from one cell to the other (see Figure 27.12); pili are sometimes referred to as *sex pili*.

Motility

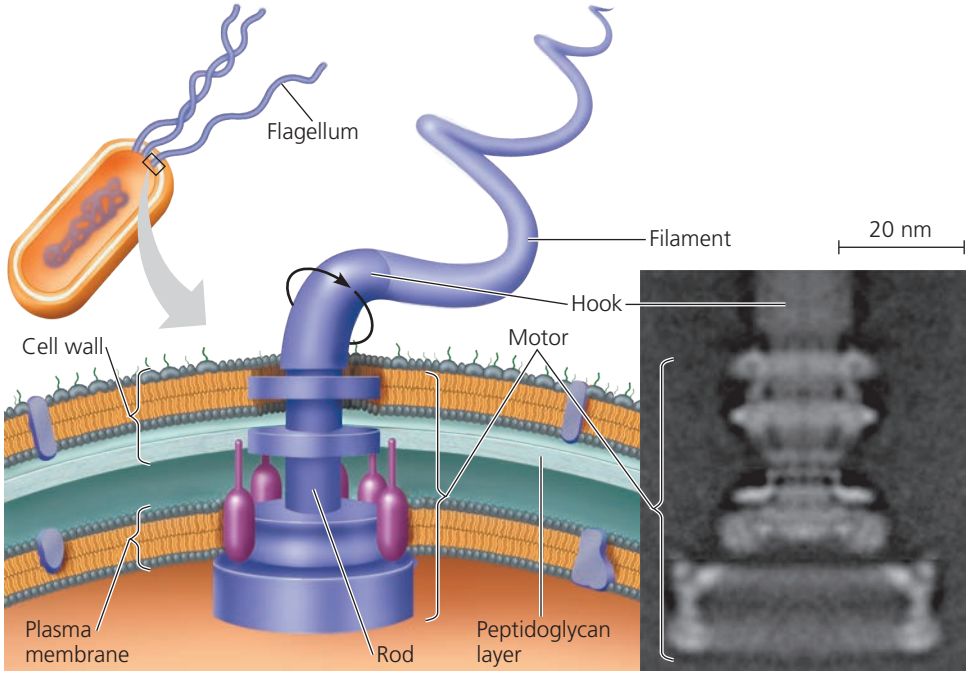
About half of all prokaryotes are capable of **taxis**, a directed movement toward or away from a stimulus (from the Greek *taxis*, to arrange). For example, prokaryotes that exhibit *chemotaxis* change their movement pattern in response to chemicals. They may move *toward* nutrients or oxygen (positive chemotaxis) or *away from* a toxic substance (negative chemotaxis). Some species can move at velocities exceeding 50 µm/sec—up to 50 times their body length per second. For perspective, consider that a person 1.7 m tall moving that fast would be running 306 km (190 miles) per hour!

Of the various structures that enable prokaryotes to move, the most common are flagella **(Figure 27.7)**. Flagella (singular, *flagellum*) may be scattered over the entire surface of the cell or concentrated at one or both ends. Prokaryotic flagella differ greatly from eukaryotic flagella: They are one-tenth the width and typically are not covered by an extension of the plasma membrane (see Figure 7.24). The flagella of prokaryotes and eukaryotes also differ in their molecular composition and their mechanism of propulsion. Among prokaryotes, bacterial and archaeal flagella are similar in size and rotational mechanism, but they are composed of entirely different and unrelated proteins. Overall, these structural and molecular comparisons indicate that the flagella of bacteria, archaea, and eukaryotes arose independently. Since current evidence shows that the flagella of organisms in the three domains perform similar functions but are not related by common descent, they are described as analogous, not homologous, structures (see Concept 21.2).

Evolutionary Origins of Bacterial Flagella

The bacterial flagellum shown in Figure 27.7 has three main parts (the motor, hook, and filament) that are themselves composed of 42 different kinds of proteins. How could such a complex structure evolve? In fact, much evidence indicates that bacterial flagella originated as simpler structures that were modified in a stepwise fashion over time. As in the case of the human eye (see Concept 25.6), biologists asked whether a less complex version of the flagellum could still benefit its owner. Analyses of hundreds of bacterial genomes indicate that only half of the flagellum's protein components appear to be necessary for it to function; the others are inessential or not encoded in the genomes of some species.

Figure 27.7 A prokaryotic flagellum. The motor of a prokaryotic flagellum consists of a system of rings embedded in the cell wall and plasma membrane (TEM). The electron transport chain pumps protons out of the cell. The diffusion of protons back into the cell provides the force that turns a curved hook and thereby causes the attached filament to rotate and propel the cell. (This diagram shows flagellar structures characteristic of gram-negative bacteria.)



VISUAL SKILLS ➤ *Predict which of the four protein rings shown in this diagram are likely hydrophobic. Explain your answer.*

Video: Prokaryotic Flagella

Of the 21 proteins required by all species studied to date, 19 are modified versions of proteins that perform other tasks in bacteria. For example, a set of 10 proteins in the motor is homologous to 10 similar proteins in a secretory system found in bacteria. (A secretory system is a protein complex that enables a cell to produce and release certain macromolecules.) Two other proteins in the motor are homologous to proteins that function in ion transport. The proteins that comprise the rod, hook, and filament are all related to each other and are descended from an ancestral protein that formed a pilus-like tube. These findings suggest that the bacterial flagellum evolved as other proteins were added to an ancestral secretory system. This is an example of *exaptation*, the process in which structures originally adapted for one function take on new functions through descent with modification.

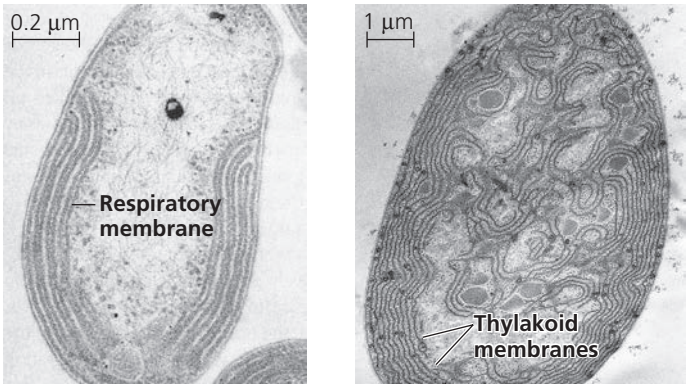
Internal Organization and DNA

The cells of prokaryotes are simpler than those of eukaryotes in both their internal structure and the physical arrangement of their DNA (see Figure 7.5). Prokaryotic cells lack the complex compartmentalization associated with the membrane-enclosed organelles found in eukaryotic cells. However, some prokaryotic cells do have specialized membranes that perform metabolic functions **(Figure 27.8)**. These membranes are usually infoldings of the plasma membrane. Recent discoveries also indicate that some prokaryotes can store metabolic by-products in simple compartments that are made out of proteins; these compartments do not have a membrane.

The genome of a prokaryote is structurally different from a eukaryotic genome and in most cases has considerably less DNA. Prokaryotes generally have circular chromosomes **(Figure 27.9)**, whereas eukaryotes have linear chromosomes. In addition, in prokaryotes the chromosome is associated with many fewer proteins than are the chromosomes of eukaryotes. Also unlike eukaryotes, prokaryotes lack a nucleus; their chromosome is located in the **nucleoid**, a region of cytoplasm that is not enclosed by a membrane. In addition to its single chromosome, a typical prokaryotic cell may also have much smaller rings of independently replicating DNA molecules called **plasmids** (see Figure 27.9), most carrying only a few genes.

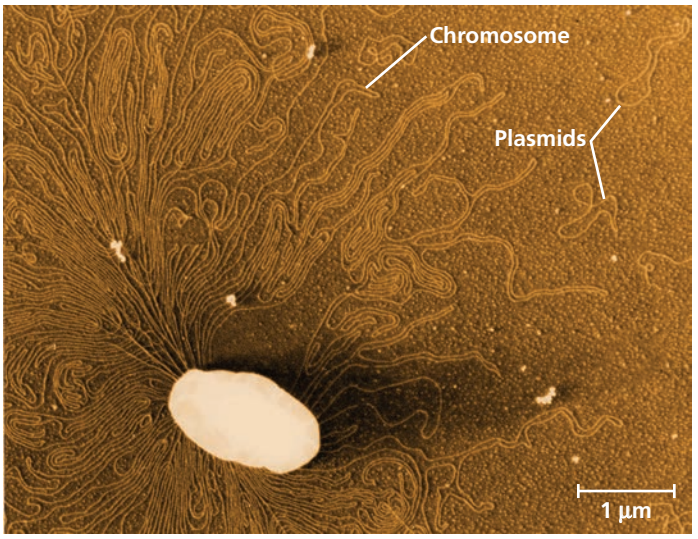
Although DNA replication, transcription, and translation are fundamentally similar processes in prokaryotes and eukaryotes, some of the details are different (see Chapter 17). For example, prokaryotic ribosomes are slightly smaller than eukaryotic ribosomes and differ in their protein and RNA content. These differences allow certain antibiotics, such as erythromycin and tetracycline, to bind to ribosomes and block protein synthesis in prokaryotes but not in eukaryotes. As a result, people can use these antibiotics to kill or inhibit the growth of bacteria without harming themselves.

Figure 27.8 Specialized membranes of prokaryotes. **(a)** Infoldings of the plasma membrane, reminiscent of the cristae of mitochondria, function in cellular respiration in some aerobic prokaryotes (TEM). **(b)** Photosynthetic prokaryotes called cyanobacteria have thylakoid membranes, much like those in chloroplasts (TEM).



(a) Aerobic prokaryote **(b) Photosynthetic prokaryote**

Figure 27.9 A prokaryotic chromosome and plasmids. The thin, tangled loops surrounding this ruptured *Escherichia coli* cell are parts of the cell's large, circular chromosome (colorized TEM). Three of the cell's plasmids, the much smaller rings of DNA, are also shown.



Reproduction

Many prokaryotes can reproduce quickly in favorable environments. By *binary fission* (see Figure 12.12), a single prokaryotic cell divides into 2 cells, which then divide into 4, 8, 16, and so on. Under optimal conditions, many prokaryotes can divide every 1–3 hours; some species can produce a new generation in only 20 minutes. At this rate, a single prokaryotic cell could give rise to a colony outweighing Earth in only two days!

In reality, of course, this does not occur. The cells eventually exhaust their nutrient supply, poison themselves with metabolic wastes, face competition from other microorganisms, or are consumed by other organisms. Still, many prokaryotic species' potential for rapid population growth

emphasizes three key features of their biology: *They are small, they reproduce by binary fission, and they often have short generation times*. As a result, prokaryotic populations can consist of many trillions of individuals—far more than populations of multicellular eukaryotes, such as plants or animals.

Animation: Structure and Reproduction of Bacteria

CONCEPT CHECK 27.1

1. Describe two adaptations that enable prokaryotes to survive in environments too harsh for other organisms.
2. Why are most of the pathogenic bacteria gram-negative?
3. **MAKE CONNECTIONS ➤** Suggest a hypothesis to explain why the thylakoid membranes of chloroplasts resemble those of cyanobacteria. Refer to Figures 7.18 and 22.21.

For suggested answers, see Appendix A.

CONCEPT 27.2

Rapid reproduction, mutation, and genetic recombination promote genetic diversity in prokaryotes

As we discussed in Unit Four, evolution cannot occur without genetic variation. The diverse adaptations exhibited by prokaryotes suggest that their populations must have considerable genetic variation—and they do. In this section, we'll examine three factors that give rise to high levels of genetic diversity in prokaryotes: rapid reproduction, mutation, and genetic recombination.

Rapid Reproduction and Mutation

In sexually reproducing species, the generation of a novel allele by a new mutation is rare for any particular gene. Instead, most of the genetic variation in sexual populations results from the way existing alleles are arranged in new combinations during meiosis and fertilization (see Concept 13.4). Prokaryotes do not reproduce sexually, so at first glance their extensive genetic variation may seem puzzling. But in many species, this variation can result from a combination of rapid reproduction and mutation.

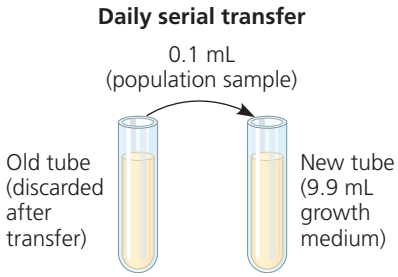
Consider the bacterium *Escherichia coli* as it reproduces by binary fission in a human intestine, one of its natural environments. After repeated rounds of division, most of the offspring cells are genetically identical to the original parent cell. However, if errors occur during DNA replication, some of the offspring cells may differ genetically. The probability of such a mutation occurring in a given *E. coli* gene is about one in 10 million (1×10^{-7}) per cell division. But among the 2×10^{10} new *E. coli* cells that arise each day in a person's intestine, there will be approximately $(2 \times 10^{10}) \times (1 \times 10^{-7}) = 2{,}000$ bacteria that have a mutation in that gene. The total number of new mutations when all 4,300 *E. coli* genes are considered is about 4,300 × 2,000—more than 8 million per day per human host.

The key point is that new mutations, though rare on a per gene basis, can increase genetic diversity quickly in species with short generation times and large populations. This diversity, in turn, can lead to rapid evolution **(Figure 27.10)**: Individuals that are genetically better equipped for their

Figure 27.10

Inquiry Can prokaryotes evolve rapidly in response to environmental change?

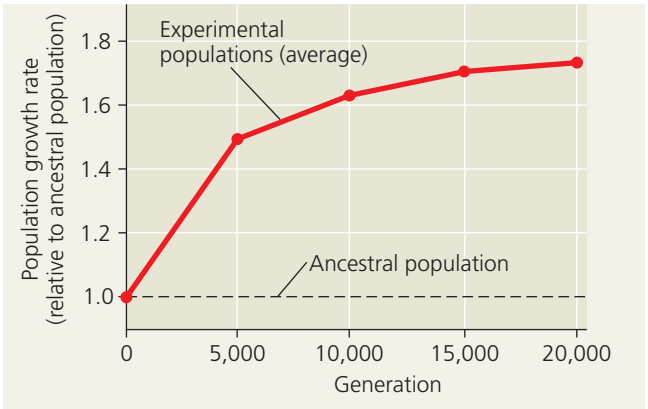
Experiment Vaughn Cooper and Richard Lenski tested the ability of *E. coli* populations to adapt to a new environment. They established 12 populations, each founded by a single cell from an *E. coli* strain, and followed these populations for 20,000 generations (3,000 days).



To maintain a continual supply of resources, each day the researchers performed a *serial transfer*: They transferred 0.1 mL of each population to a new tube containing 9.9 mL of fresh growth medium. The growth medium used throughout the experiment provided a challenging environment that contained only low levels of glucose and other resources needed for growth.

Samples were periodically removed from the 12 populations and grown in competition with the common ancestral strain in the experimental (low-glucose) environment.

Results The fitness of the experimental populations, as measured by the growth rate of each population, increased rapidly for the first 5,000 generations (2 years) and more slowly for the next 15,000 generations. The graph shows the averages for the 12 populations.



Conclusion Populations of *E. coli* continued to accumulate beneficial mutations for 20,000 generations, allowing rapid evolution of increased population growth rates in their new environment.

Data from V. S. Cooper and R. E. Lenski, The population genetics of ecological specialization in evolving *Escherichia coli* populations, *Nature* 407:736–739 (2000).

WHAT IF? ➤ *Suggest possible functions of the genes whose sequence or expression was altered as the experimental populations evolved in the low-glucose environment.*

environment tend to survive and reproduce at higher rates than other individuals. The ability of prokaryotes to adapt rapidly to new conditions highlights the point that although the structure of their cells is simpler than that of eukaryotic cells, prokaryotes are not "primitive" or "inferior" in an evolutionary sense. They are, in fact, highly evolved: For 3.5 billion years, prokaryotic populations have responded successfully to many types of environmental challenges.

Genetic Recombination

Although new mutations are a major source of variation in prokaryotic populations, additional diversity arises from *genetic recombination*, the combining of DNA from two sources. In eukaryotes, the sexual processes of meiosis and fertilization combine DNA from two individuals in a single zygote. But meiosis and fertilization do not occur in prokaryotes. Instead, three other mechanisms—transformation, transduction, and conjugation—can bring together prokaryotic DNA from different individuals (that is, different cells). When the individuals are members of different species, this movement of genes from one organism to another is called *horizontal gene transfer*. Although scientists have found evidence that each of these mechanisms can transfer DNA within and between species in both domain Bacteria and domain Archaea, to date most of our knowledge comes from research on bacteria.

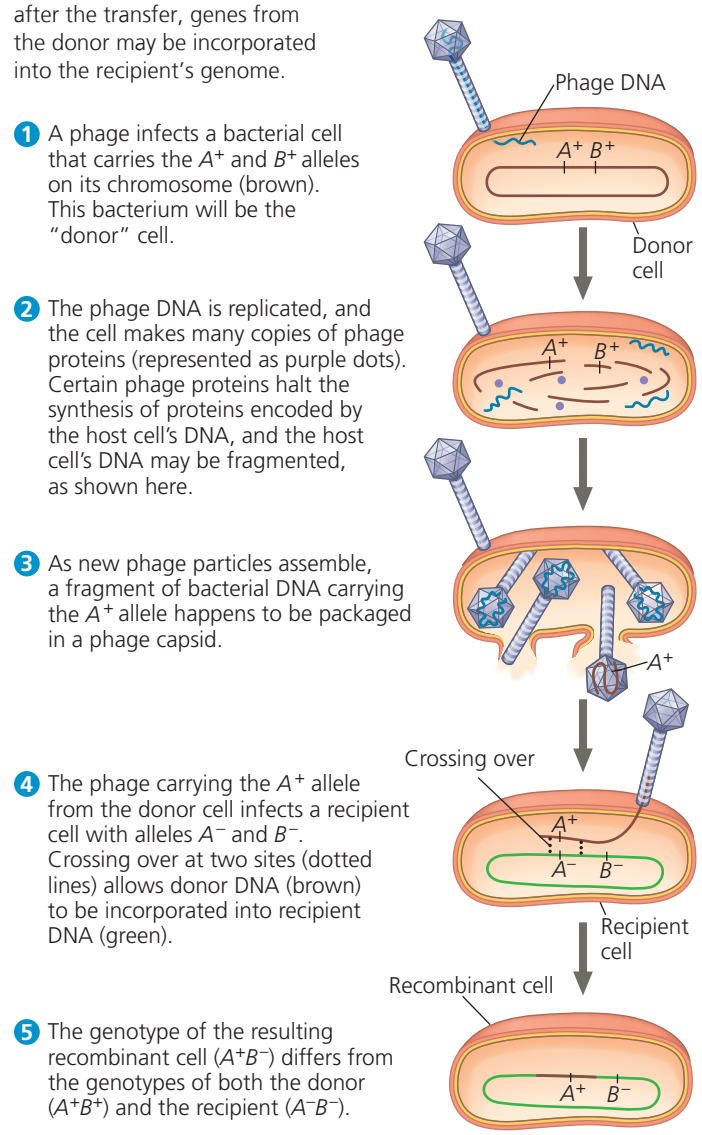
Transformation and Transduction

In **transformation**, the genotype and possibly phenotype of a prokaryotic cell are altered by the uptake of foreign DNA from its surroundings. For example, a harmless strain of *Streptococcus pneumoniae* can be transformed into pneumonia-causing cells if the cells are exposed to DNA from a pathogenic strain (see Concept 16.1). This transformation occurs when a nonpathogenic cell takes up a piece of DNA carrying the allele for pathogenicity and replaces its own allele with the foreign allele, an exchange of homologous DNA segments. The cell is now a recombinant: Its chromosome contains DNA derived from two different cells.

For many years after transformation was discovered in laboratory cultures, most biologists thought it was too rare and haphazard to play an important role in natural bacterial populations. But researchers have since learned that many bacteria have cell-surface proteins that recognize DNA from closely related species and transport it into the cell. Once inside the cell, the foreign DNA can be incorporated into the genome by homologous DNA exchange.

In **transduction**, phages (from "bacteriophages," the viruses that infect bacteria) carry prokaryotic genes from one host cell to another. In most cases, transduction results from accidents that occur during the phage replicative cycle **(Figure 27.11)**. A virus that carries prokaryotic DNA may not be able to replicate because it lacks some or all of its own genetic material. However, the virus can attach to another prokaryotic cell (a recipient) and inject prokaryotic DNA acquired from the first cell (the donor). If some of this DNA is then incorporated into the recipient cell's chromosome by crossing over, a recombinant cell is formed.

Figure 27.11 Transduction. Phages may carry pieces of a bacterial chromosome from one cell (the donor) to another (the recipient). If crossing over occurs after the transfer, genes from the donor may be incorporated into the recipient's genome.

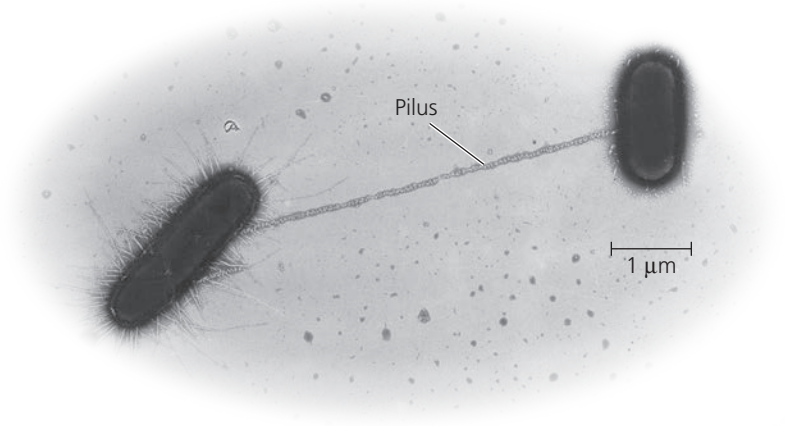


VISUAL SKILLS ➤ *Based on this diagram, describe the circumstances in which a transduction event would result in horizontal gene transfer.*

Conjugation and Plasmids

In a process called **conjugation**, DNA is transferred between two prokaryotic cells (usually of the same species) that are temporarily joined. In bacteria, the DNA transfer is always one-way: One cell donates the DNA, and the other receives it. We'll focus here on the mechanism used by *E. coli*.

Figure 27.12 Bacterial conjugation. The *E. coli* donor cell (left) extends a pilus that attaches to a recipient cell, a key first step in the transfer of DNA. The pilus is a flexible tube of protein subunits (TEM).



First, a pilus of the donor cell attaches to the recipient **(Figure 27.12)**. The pilus then retracts, pulling the two cells together, like a grappling hook. The next step is thought to be the formation of a temporary structure between the two cells, a "mating bridge" through which the donor may transfer DNA to the recipient. However, the mechanism by which DNA transfer occurs is unclear; indeed, recent evidence indicates that DNA may pass directly through the hollow pilus.

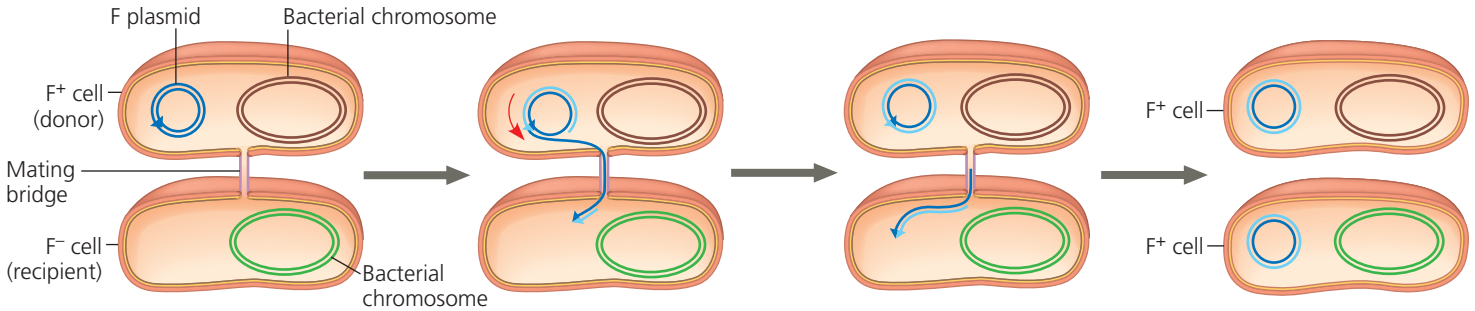
The ability to form pili and donate DNA during conjugation results from the presence of a particular piece of DNA called the **F factor** (F for *f*ertility). The F factor of *E. coli* consists of about 25 genes, most required for the production of pili. As shown in **Figure 27.13**, the F factor can exist either as a plasmid or as a segment of DNA within the bacterial chromosome.

The F Factor as a Plasmid

The F factor in its plasmid form is called the **F plasmid**. Cells containing the F plasmid,

Figure 27.13 Conjugation and recombination in *E. coli*. The DNA replication that accompanies transfer of an F plasmid or part of an Hfr bacterial chromosome is called *rolling circle replication*. In effect, the intact circular parental DNA strand "rolls" as its other strand peels off and a new complementary strand is synthesized.

MB Figure Walkthrough



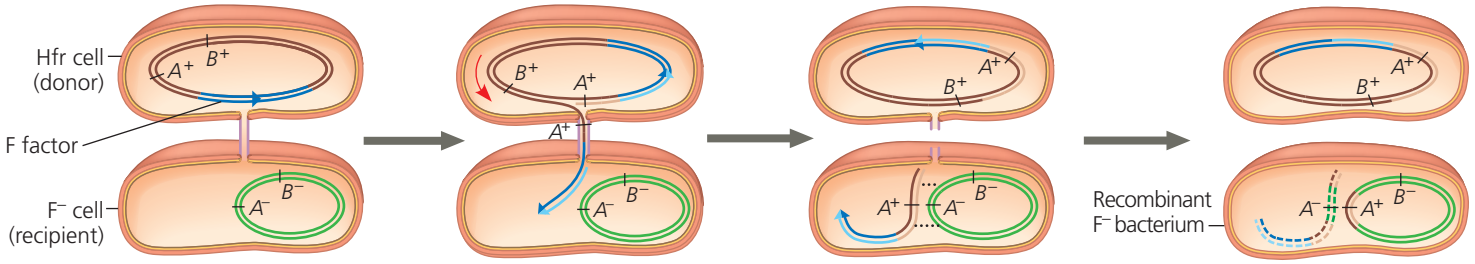
1 A cell carrying an F plasmid (an F^+ cell) forms a mating bridge with an F^- cell. One strand of the plasmid's DNA breaks at the point marked by the arrowhead.

2 Using the unbroken strand as a template, the cell synthesizes a new strand (light blue). Meanwhile, the broken strand peels off (red arrow), and one end enters the F^- cell. There synthesis of its complementary strand begins.

3 DNA replication continues in both the donor and recipient cells, as the transferred plasmid strand moves farther into the recipient cell.

4 Once DNA transfer and synthesis are completed, the plasmid in the recipient cell circularizes. The recipient cell is now a recombinant F^+ cell.

(a) Conjugation and transfer of an F plasmid



1 In an Hfr cell, the F factor (dark blue) is integrated into the bacterial chromosome. Since an Hfr cell has all of the F factor genes, it can form a mating bridge with an F^- cell and transfer DNA.

2 A single strand of the F factor breaks and begins to move through the bridge. DNA replication occurs in both donor and recipient cells, resulting in double-stranded DNA (daughter strands shown in lighter color).

3 The mating bridge usually breaks before the entire chromosome is transferred. Crossing over at two sites (dotted lines) can result in the exchange of homologous genes (here, A^+ and A^-) between the transferred DNA (brown) and the recipient's chromosome (green).

4 Cellular enzymes degrade any linear DNA not incorporated into the chromosome. The recipient cell, with a new combination of genes but no F factor, is now a recombinant F^- cell.

(b) Conjugation and transfer of part of an Hfr bacterial chromosome, resulting in recombination. A^+/A^- and B^+/B^- indicate alleles for gene *A* and gene *B*, respectively.

designated F^+ cells, function as DNA donors during conjugation **(Figure 27.13a)**. Cells lacking the F factor, designated F^-, function as DNA recipients during conjugation. The F^+ condition is transferable in the sense that an F^+ cell converts an F^- cell to F^+ if a copy of the entire F plasmid is transferred. Even if this does not occur, as long as some of the F plasmid's DNA is transferred successfully to the recipient cell, that cell is now a recombinant cell.

The F Factor in the Chromosome Chromosomal genes can be transferred during conjugation when the donor cell's F factor is integrated into the chromosome. A cell with the F factor built into its chromosome is called an *Hfr cell* (for *h*igh *f*requency of *r*ecombination). Like an F^+ cell, an Hfr cell functions as a donor during conjugation with an F^- cell **(Figure 27.13b)**. When chromosomal DNA from an Hfr cell enters an F^- cell, homologous regions of the Hfr and F^- chromosomes may align, allowing segments of their DNA to be exchanged. As a result, the recipient cell becomes a recombinant bacterium that has genes derived from the chromosomes of two different cells—a new genetic variant on which evolution can act.

R Plasmids and Antibiotic Resistance During the 1950s in Japan, physicians started noticing that some hospital patients with bacterial dysentery, which produces severe diarrhea, did not respond to antibiotics that had been effective in the past. Apparently, resistance to these antibiotics had evolved in some strains of *Shigella*, the bacterium that causes the disease.

Eventually, researchers began to identify the specific genes that confer antibiotic resistance in *Shigella* and other pathogenic bacteria. Sometimes mutation in a chromosomal gene of the pathogen can confer resistance. For example, a mutation in one gene may make it less likely that the pathogen will transport a particular antibiotic into its cell. Mutation in a different gene may alter the intracellular target protein for an antibiotic molecule, reducing its inhibitory effect. In other cases, bacteria have "resistance genes," which code for enzymes that specifically destroy or otherwise hinder the effectiveness of certain antibiotics, such as tetracycline or ampicillin. Such resistance genes are often carried by plasmids known as **R plasmids** (R for *r*esistance).

Exposing a bacterial population to a specific antibiotic will kill antibiotic-sensitive bacteria but not those that happen to have R plasmids with genes that counter the antibiotic. Under these circumstances, we would predict that natural selection would cause the fraction of the bacterial population carrying genes for antibiotic resistance to increase, and that is exactly what happens. The medical consequences are also predictable: Resistant strains of pathogens are becoming more common, making the treatment of certain bacterial infections more difficult. The problem is compounded by the fact that many R plasmids, like F plasmids, have genes that encode pili and enable DNA transfer from one bacterial cell to another by conjugation. Making the problem still worse, some R plasmids carry genes for resistance to as many as ten antibiotics.

CONCEPT CHECK 27.2

1. Although rare on a per gene basis, new mutations can add considerable genetic variation to prokaryotic populations in each generation. Explain how this occurs.
2. Distinguish between the three mechanisms by which bacteria can transfer DNA from one bacterial cell to another.
3. Why do F^- cells always serve as DNA recipients during the process of conjugation between two prokaryotic cells?
4. **WHAT IF? ➤** If a nonpathogenic bacterium were to acquire resistance to antibiotics, could this strain pose a health risk to people? In general, how does DNA transfer among bacteria affect the spread of resistance genes?

For suggested answers, see Appendix A.

CONCEPT 27.3

Diverse nutritional and metabolic adaptations have evolved in prokaryotes

The extensive genetic variation found in prokaryotes is reflected in their diverse nutritional adaptations. Like all organisms, prokaryotes can be categorized by how they obtain energy and the carbon used in building the organic molecules that make up cells. Every type of nutrition observed in eukaryotes is represented among prokaryotes, along with some nutritional modes unique to prokaryotes. In fact, prokaryotes have an astounding range of metabolic adaptations, much broader than that found in eukaryotes.

Organisms that obtain energy from light are called *phototrophs*, and those that obtain energy from chemicals are called *chemotrophs*. Organisms that need only CO_2 or related compounds as a carbon source are called *autotrophs*. In contrast, *heterotrophs* require at least one organic nutrient, such as glucose, to make other organic compounds. Combining possible energy sources and carbon sources results in four major modes of nutrition, summarized in **Table 27.1**.

The Role of Oxygen in Metabolism

Prokaryotic metabolism also varies with respect to oxygen (O_2). **Obligate aerobes** must use O_2 for cellular respiration and cannot grow without it. **Obligate anaerobes**, on the other hand, are poisoned by O_2. Some obligate anaerobes live exclusively by fermentation; others extract chemical energy by **anaerobic respiration**, in which substances other than O_2, such as nitrate ions (NO_3^-) or sulfate ions (SO_4^{2-}), accept electrons at the "downhill" end of electron transport chains. **Facultative anaerobes** use O_2 if it is present but can also carry out fermentation or anaerobic respiration in an anaerobic environment.

Nitrogen Metabolism

Nitrogen is essential for the production of amino acids and nucleic acids in all organisms. Whereas eukaryotes can obtain

Table 27.1 Major Nutritional Modes

Mode	Energy Source	Carbon Source	Types of Organisms
AUTOTROPH			
Photoautotroph	Light	CO_2, HCO_3^-, or related compound	Photosynthetic prokaryotes (for example, cyanobacteria); plants; certain protists (for example, algae)
Chemoautotroph	Inorganic chemicals (such as H_2S, NH_3, or Fe^{2+})	CO_2, HCO_3^-, or related compound	Unique to certain prokaryotes (for example, *Sulfolobus*)
HETEROTROPH			
Photoheterotroph	Light	Organic compounds	Unique to certain aquatic and salt-loving prokaryotes (for example, *Rhodobacter*, *Chloroflexus*)
Chemoheterotroph	Organic compounds	Organic compounds	Many prokaryotes (for example, *Clostridium*) and protists; fungi; animals; some plants

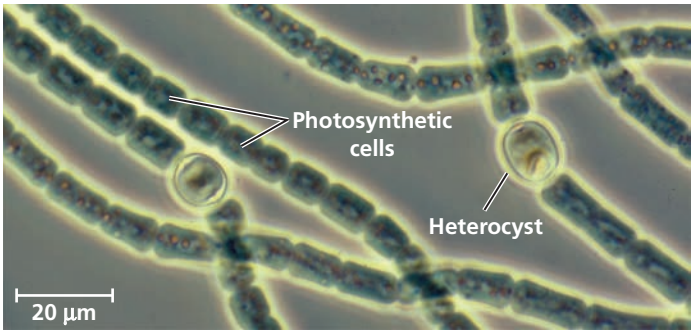
nitrogen only from a limited group of nitrogen compounds, prokaryotes can metabolize nitrogen in many forms. For example, some cyanobacteria and some methanogens (a group of archaea) convert atmospheric nitrogen (N_2) to ammonia (NH_3), a process called **nitrogen fixation**. The cells can then incorporate this "fixed" nitrogen into amino acids and other organic molecules. In terms of nutrition, nitrogen-fixing cyanobacteria are some of the most self-sufficient organisms, since they need only light, CO_2, N_2, water, and some minerals to grow.

Nitrogen fixation has a large impact on other organisms. For example, nitrogen-fixing prokaryotes can increase the nitrogen available to plants, which cannot use atmospheric nitrogen but can use the nitrogen compounds that the prokaryotes produce from ammonia. Concept 55.4 discusses this and other essential roles that prokaryotes play in the nitrogen cycles of ecosystems.

Metabolic Cooperation

Cooperation between prokaryotic cells allows them to use environmental resources they could not use as individual cells. In some cases, this cooperation takes place between specialized cells of a filament. For instance, the cyanobacterium *Anabaena* has genes that encode proteins for photosynthesis and for nitrogen fixation. However, a single cell cannot carry out both processes at the same time because photosynthesis produces O_2, which inactivates the enzymes involved in nitrogen fixation. Instead of living as isolated cells, *Anabaena* forms filamentous chains **(Figure 27.14)**. Most cells in a filament carry out only photosynthesis, while a few specialized cells called **heterocysts** (sometimes called *heterocytes*) carry out only nitrogen fixation. Each heterocyst is surrounded by a thickened cell wall that restricts entry of O_2 produced by neighboring photosynthetic cells. Intercellular connections allow heterocysts to transport fixed nitrogen to neighboring cells and to receive carbohydrates.

Figure 27.14 Metabolic cooperation in a prokaryote. In the filamentous freshwater cyanobacterium *Anabaena*, heterocysts fix nitrogen, while the other cells carry out photosynthesis (LM).



Metabolic cooperation between different prokaryotic species often occurs in surface-coating colonies known as **biofilms**. Cells in a biofilm secrete signaling molecules that recruit nearby cells, causing the colonies to grow. The cells also produce polysaccharides and proteins that stick the cells to the substrate and to one another; these polysaccharides and proteins form the capsule, or slime layer, mentioned earlier in the chapter. Channels in the biofilm allow nutrients to reach cells in the interior and wastes to be expelled. Biofilms are common in nature, but they can cause problems by contaminating industrial products and medical equipment and contributing to tooth decay and more serious health problems. Altogether, damage caused by biofilms costs billions of dollars annually.

In another example of cooperation between prokaryotes, sulfate-consuming bacteria coexist with methane-consuming archaea in ball-shaped aggregates on the ocean floor. The bacteria appear to use the archaea's waste products, such as organic compounds and hydrogen. In turn, the bacteria produce sulfur compounds that the archaea use as oxidizing agents when they consume methane in the absence of oxygen. This partnership has global ramifications: Each year, these archaea consume an estimated 300 billion kg of methane, a major greenhouse gas (see Concept 56.4).

CONCEPT CHECK 27.3

1. Distinguish between the four major modes of nutrition, noting which are unique to prokaryotes.
2. A bacterium requires only the amino acid methionine as an organic nutrient and lives in lightless caves. What mode of nutrition does it employ? Explain.
3. **WHAT IF? ➤** Describe what you might eat for a typical meal if humans, like cyanobacteria, could fix nitrogen.

For suggested answers, see Appendix A.

CONCEPT 27.4

Prokaryotes have radiated into a diverse set of lineages

Since their origin 3.5 billion years ago, prokaryotic populations have radiated extensively as a wide range of structural and metabolic adaptations have evolved in them. Collectively, these adaptations have enabled prokaryotes to inhabit every environment known to support life—if there are organisms in a particular place, some of those organisms are prokaryotes. Yet despite their obvious success, it is only in recent decades that advances in genomics have begun to reveal the full extent of prokaryotic diversity.

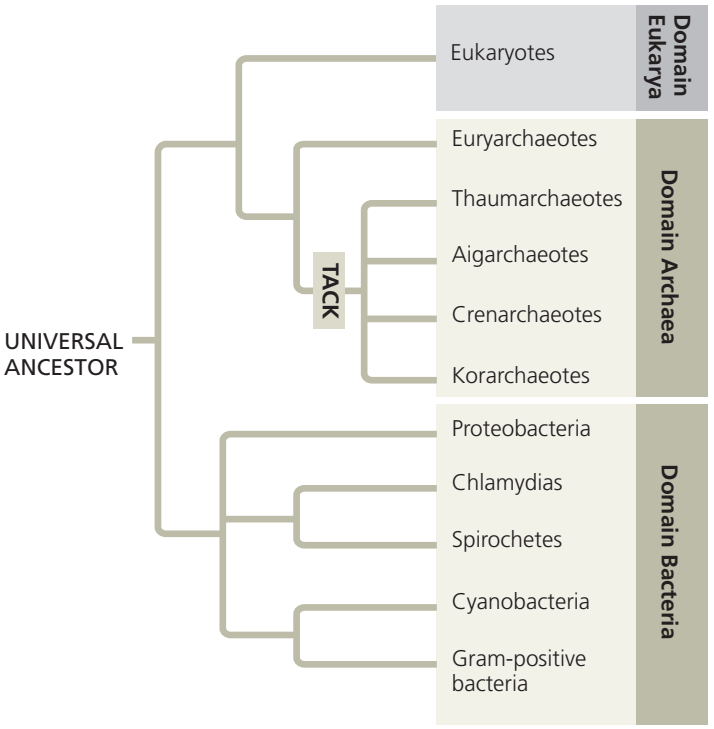
An Overview of Prokaryotic Diversity

In the 1970s, microbiologists began using small-subunit ribosomal RNA as a marker for evolutionary relationships. Their results indicated that many prokaryotes once classified as bacteria are actually more closely related to eukaryotes and belong in a domain of their own: Archaea. Microbiologists have since analyzed larger amounts of genetic data—including more than 1,700 entire genomes—and have concluded that a few traditional taxonomic groups, such as cyanobacteria, are monophyletic. However, other traditional groups, such as gram-negative bacteria, are scattered throughout several lineages. **Figure 27.15** shows one phylogenetic hypothesis for some of the major taxa of prokaryotes based on molecular systematics.

One lesson from studying prokaryotic phylogeny is that the genetic diversity of prokaryotes is immense. When researchers began to sequence the genes of prokaryotes, they could investigate only the small fraction of species that could be cultured in the laboratory. In the 1980s, researchers began using the polymerase chain reaction (PCR; see Figure 19.8) to analyze the genes of prokaryotes collected from the environment (such as from soil or water samples). Such "genetic prospecting" is now widely used; in fact, today entire prokaryotic genomes can be obtained from environmental samples using *metagenomics* (see Concept 20.1). Each year these techniques add new branches to the tree of life. While only about 10,600 prokaryotic species worldwide have been assigned scientific names, a single handful of soil could contain 10,000 prokaryotic species by some estimates. Taking full stock of this diversity will require many years of research.

Another important lesson from molecular systematics is that horizontal gene transfer has played a key role in the evolution of prokaryotes. Over hundreds of millions of years, prokaryotes have acquired genes from even distantly related species, and they continue to do so today. As a result, significant portions of the genomes of many prokaryotes are actually mosaics of genes imported from other species. For example, a study of 329 sequenced bacterial genomes found that an average of 75% of the genes in each genome had been transferred horizontally at some point in their evolutionary history. As we saw in Concept 22.6, such gene transfers can make it difficult to determine phylogenetic relationships. Still, it is clear that for billions of years, the prokaryotes have evolved in two separate lineages, the bacteria and the archaea (see Figure 27.15).

Figure 27.15 A simplified phylogeny of prokaryotes. This tree shows relationships among major prokaryotic groups based on molecular data; some of these relationships are shown as polytomies to reflect their uncertain order of divergence. Recent studies indicate that within Archaea, the thaumarchaeotes, aigarchaeotes, crenarchaeotes, and korarchaeotes are closely related; systematists have placed them in a supergroup called "TACK" in reference to the first letters of their names.



VISUAL SKILLS *Based on this phylogenetic tree diagram, which domain is the sister group of Archaea?*

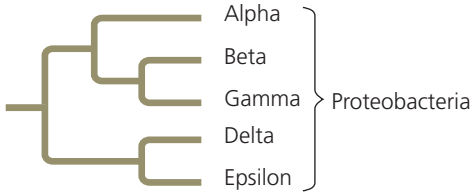
Bacteria

As surveyed in **Figure 27.16**, bacteria include the vast majority of prokaryotic species familiar to most people, from the pathogenic species that cause strep throat and tuberculosis to the beneficial species used to make Swiss cheese and yogurt. Every major mode of nutrition and metabolism is represented among bacteria, and even a small taxonomic group of bacteria may contain species exhibiting many different nutritional modes. As we'll see, the diverse nutritional and metabolic capabilities of bacteria—and archaea—are behind the great impact these organisms have on Earth and its life.

Figure 27.16 Exploring Selected Major Groups of Bacteria

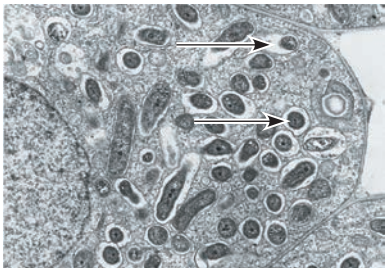
Proteobacteria

This large and diverse clade of gram-negative bacteria includes photoautotrophs, chemoautotrophs, and heterotrophs. Some are anaerobic, while others are aerobic. Molecular systematists currently recognize five subgroups of proteobacteria; the phylogenetic tree at right shows their relationships based on molecular data.



Subgroup: Alpha Proteobacteria

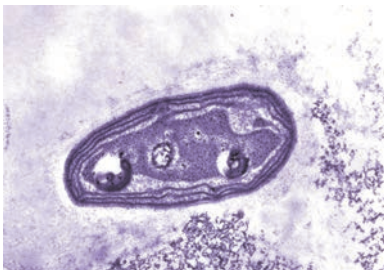
Many of the species in this subgroup are closely associated with eukaryotic hosts. For example, *Rhizobium* species live in nodules within the roots of legumes (plants of the pea/bean family), where the bacteria convert atmospheric N_2 to compounds the host plant can use to make proteins. Species in the genus *Agrobacterium* produce tumors in plants; genetic engineers use these bacteria to carry foreign DNA into the genomes of crop plants. Scientists hypothesize that mitochondria evolved from aerobic alpha proteobacteria through endosymbiosis.



Rhizobium (arrows) inside a root cell of a legume (TEM)

Subgroup: Beta Proteobacteria

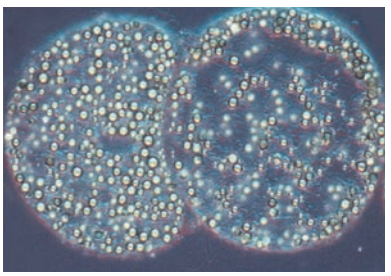
This nutritionally diverse subgroup includes *Nitrosomonas*, a genus of soil bacteria that play an important role in nitrogen recycling by oxidizing ammonium (NH_4^+), producing nitrite (NO_2^-) as a waste product. Other members of this subgroup include a wide range of aquatic species, such as the photoheterotroph *Rubrivivax*, along with pathogens such as the species that causes the sexually transmitted disease gonorrhea, *Neisseria gonorrhoeae*.



Nitrosomonas (colorized TEM)

Subgroup: Gamma Proteobacteria

This subgroup's autotrophic members include sulfur bacteria, such as *Thiomargarita namibiensis*, which obtain energy by oxidizing H_2S, producing sulfur as a waste product (the small globules in the photograph at right). Some heterotrophic gamma proteobacteria are pathogens; for example, *Legionella* causes Legionnaires' disease, *Salmonella* is responsible for some cases of food poisoning, and *Vibrio cholerae* causes cholera. *Escherichia coli*, a common resident of the intestines of humans and other mammals, normally is not pathogenic.



Thiomargarita namibiensis containing sulfur wastes (LM)

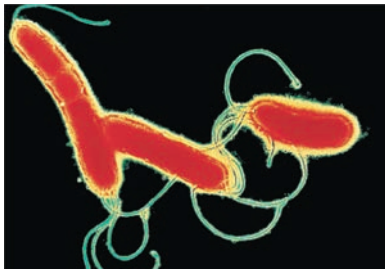
Subgroup: Delta Proteobacteria

This subgroup includes the slime-secreting myxobacteria. When the soil dries out or food is scarce, the cells congregate into a fruiting body that releases resistant "myxospores." These cells found new colonies in favorable environments. Another group of delta proteobacteria, the bdellovibrios, attack other bacteria, charging at up to 100 µm/sec (comparable to a human running 240 km/hr). The attack begins when a bdellovibrio attaches to specific molecules found on the outer covering of some bacterial species. The bdellovibrio then drills into its prey by using digestive enzymes and spinning at 100 revolutions per second.

Fruiting bodies of *Chondromyces crocatus*, a myxobacterium (SEM)

Subgroup: Epsilon Proteobacteria

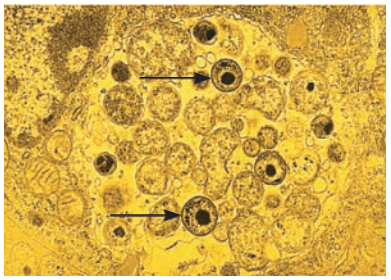
Most species in this subgroup are pathogenic to humans or other animals. Epsilon proteobacteria include *Campylobacter*, which causes blood poisoning and intestinal inflammation, and *Helicobacter pylori*, which causes stomach ulcers.



Helicobacter pylori (colorized TEM)

Chlamydias

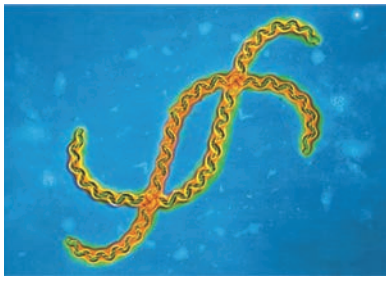
These parasites can survive only within animal cells, depending on their hosts for resources as basic as ATP. The gram-negative walls of chlamydias are unusual in that they lack peptidoglycan. One species, *Chlamydia trachomatis*, is the most common cause of blindness in the world and also causes nongonococcal urethritis, the most common sexually transmitted disease in the United States.



Chlamydia (arrows) inside an animal cell (colorized TEM)

Spirochetes

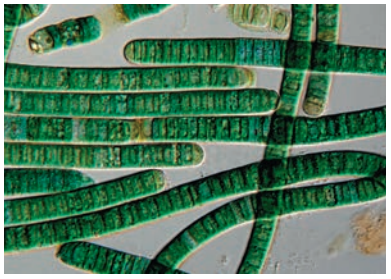
These helical gram-negative heterotrophs spiral through their environment by means of rotating, internal, flagellum-like filaments. Many spirochetes are free-living, but others are notorious pathogenic parasites: *Treponema pallidum* causes syphilis, and *Borrelia burgdorferi* causes Lyme disease.



Leptospira, a spirochete (colorized TEM)

Cyanobacteria

These gram-negative photoautotrophs are the only prokaryotes with plantlike, oxygen-generating photosynthesis. (In fact, chloroplasts are thought to have evolved from an endosymbiotic cyanobacterium.) Both solitary and filamentous cyanobacteria are abundant components of freshwater and marine *phytoplankton*, the collection of photosynthetic organisms that drift near the water's surface. Some filaments have cells specialized for nitrogen fixation, the process that incorporates atmospheric N_2 into inorganic compounds that can be used in the synthesis of amino acids and other organic molecules.



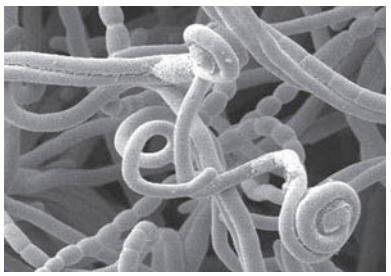
Oscillatoria, a filamentous cyanobacterium

Gram-Positive Bacteria

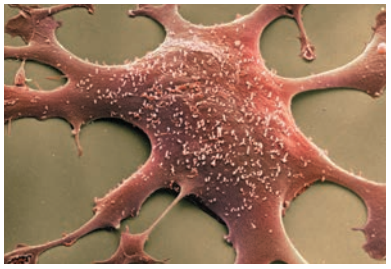
Gram-positive bacteria rival the proteobacteria in diversity. Species in one subgroup, the actinomycetes (from the Greek *mykes*, fungus, for which these bacteria were once mistaken), form colonies containing branched chains of cells. Two species of actinomycetes cause tuberculosis and leprosy. However, most actinomycetes are free-living species that help decompose the organic matter in soil; their secretions are partly responsible for the "earthy" odor of rich soil. Soil-dwelling species in the genus *Streptomyces* (top) are cultured by pharmaceutical companies as a source of many antibiotics, including streptomycin.

Gram-positive bacteria include many solitary species, such as *Bacillus anthracis*, which causes anthrax, and *Clostridium botulinum*, which causes botulism. The various species of *Staphylococcus* and *Streptococcus* are also gram-positive bacteria.

Mycoplasmas (bottom) are the only bacteria known to lack cell walls. They are also the tiniest known cells, with diameters as small as 0.1 µm, only about five times as large as a ribosome. Mycoplasmas have small genomes—*Mycoplasma genitalium* has only 517 genes, for example. Many mycoplasmas are free-living soil bacteria, but others are pathogens.



Streptomyces, the source of many antibiotics (SEM)



Hundreds of mycoplasmas covering a human fibroblast cell (colorized SEM)

Table 27.2 A Comparison of the Three Domains of Life

CHARACTERISTIC	DOMAIN		
	Bacteria	Archaea	Eukarya
Nuclear envelope	Absent	Absent	Present
Membrane-enclosed organelles	Absent	Absent	Present
Peptidoglycan in cell wall	Present	Absent	Absent
Membrane lipids	Unbranched hydrocarbons	Some branched hydrocarbons	Unbranched hydrocarbons
RNA polymerase	One kind	Several kinds	Several kinds
Initiator amino acid for protein synthesis	Formyl-methionine	Methionine	Methionine
Introns in genes	Very rare	Present in some genes	Present in many genes
Response to the antibiotics streptomycin and chloramphenicol	Growth usually inhibited	Growth not inhibited	Growth not inhibited
Histones associated with DNA	Absent	Present in some species	Present
Circular chromosome	Present	Present	Absent
Growth at temperatures > 100°C	No	Some species	No

Archaea

Archaea share certain traits with bacteria and other traits with eukaryotes **(Table 27.2)**. However, archaea also have many unique characteristics, as we would expect in a taxon that has followed a separate evolutionary path for so long.

The first prokaryotes assigned to domain Archaea live in environments so extreme that few other organisms can survive there. Such organisms are called **extremophiles**, meaning "lovers" of extreme conditions (from the Greek *philos*, lover), and include extreme halophiles and extreme thermophiles.

Extreme halophiles (from the Greek *halo*, salt) live in highly saline environments, such as the Great Salt Lake in Utah, the Dead Sea in Israel, and the Spanish lake shown in Figure 27.1. Some species merely tolerate salinity, while others require an environment that is several times saltier than seawater (which has a salinity of 3.5%). For example, the proteins and cell wall of *Halobacterium* have unusual features that improve function in extremely salty environments but render these organisms incapable of survival if the salinity drops below 9%.

Figure 27.17 Extreme thermophiles. Thermophilic prokaryotes grow in the hot water of geysers in the Valley of Geysers in Kamchatka, Russia.

MAKE CONNECTIONS ➤ *How might the enzymes of thermophiles differ from those of other organisms? (Review enzymes in Concept 6.4.)*

Extreme thermophiles (from the Greek *thermos*, hot) thrive in very hot environments **(Figure 27.17)**. For example, archaea in the genus *Sulfolobus* live in sulfur-rich volcanic springs as hot as 90°C. At temperatures this high, the cells of most organisms die because their DNA does not remain in a double helix and many of their proteins denature. *Sulfolobus* and other extreme thermophiles avoid this fate because they have structural and biochemical adaptations that make their DNA and proteins stable at high temperatures. One extreme thermophile that lives near deep-sea hot springs called *hydrothermal vents* is informally known as "strain 121," since it can reproduce even at 121°C. Another extreme thermophile, *Pyrococcus furiosus*, is used in biotechnology as a source of DNA polymerase for the PCR technique (see Figure 19.8).

Many other archaea live in more moderate environments. Consider the **methanogens**, archaea that release methane as a by-product of their unique ways of obtaining energy. Many methanogens use CO_2 to oxidize H_2, a process that produces both energy and methane waste. Among the strictest of anaerobes, methanogens are poisoned by O_2. Although some methanogens live in extreme environments, such as under kilometers of ice in Greenland, others live in swamps and marshes where other microorganisms have consumed all the O_2. The "marsh gas" found in such environments is the methane released by these archaea. Other species inhabit the anaerobic guts of cattle, termites, and other herbivores, playing an essential role in the nutrition of these animals. Methanogens are also useful to humans as decomposers in sewage treatment facilities.

Many extreme halophiles and all known methanogens are archaea in the clade Euryarchaeota (from the Greek *eurys*, broad, a reference to their wide habitat range). The

euryarchaeotes also include some extreme thermophiles, though most thermophilic species belong to a second clade, Crenarchaeota (*cren* means "spring," such as a hydrothermal spring). Metagenomic studies have identified many species of euryarchaeotes and crenarchaeotes that are not extremophiles. These archaea exist in habitats ranging from farm soils to lake sediments to the surface of the open ocean.

New findings continue to inform our understanding of archaeal phylogeny. For example, recent metagenomic studies have uncovered the genomes of many species that are not members of Euryarchaeota or Crenarchaeota. Moreover, phylogenomic analyses show that three of these newly discovered groups—the Thaumarchaeota, Aigarchaeota, and Korarchaeota—are more closely related to the Crenarchaeota than they are to the Euryarchaeota. These findings have led to the identification of a "supergroup" that contains the Thaumarchaeota, Aigarchaeota, Crenarchaeota, and Korarchaeota (see Figure 27.15). This supergroup is referred to as "TACK" based on the names of the groups it includes. In 2015, the importance of the TACK supergroup was highlighted by the discovery of the lokiarchaeotes, a group that is closely related to TACK archaea and that could possibly represent the long sought-after sister group of the eukaryotes. As such, the characteristics of lokiarchaeotes may shed light on one of the major puzzles of biology today—how eukaryotes arose from their prokaryotic ancestors. The pace of these and other recent discoveries suggests that as metagenomic prospecting continues, the tree in Figure 27.15 will likely undergo further changes.

CONCEPT CHECK 27.4

1. Explain how molecular systematics and metagenomics have contributed to our understanding of the phylogeny and evolution of prokaryotes.
2. **WHAT IF?** ➤ What would the discovery of a bacterial species that is a methanogen imply about the evolution of the methane-producing pathway?

For suggested answers, see Appendix A.

CONCEPT 27.5

Prokaryotes play crucial roles in the biosphere

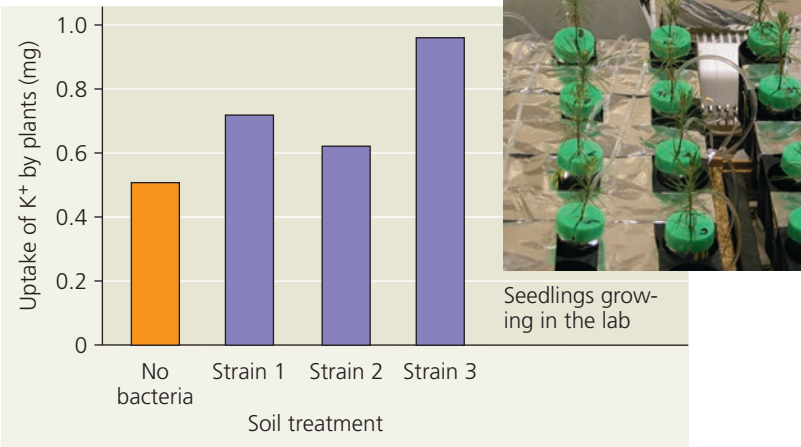
If people were to disappear from the planet tomorrow, life on Earth would change for many species, but few would be driven to extinction. In contrast, prokaryotes are so important to the biosphere that if they were to disappear, the prospects of survival for many other species would be dim.

Chemical Recycling

The atoms that make up the organic molecules in all living things were at one time part of inorganic substances in the soil, air, and water. Sooner or later, those atoms will return to the nonliving environment. Ecosystems depend on the continual recycling of chemical elements between the living and nonliving components of the environment, and prokaryotes play a major role in this process. For example, some chemoheterotrophic prokaryotes function as **decomposers**, breaking down dead organisms as well as waste products and thereby unlocking supplies of carbon, nitrogen, and other elements. Without the actions of prokaryotes and other decomposers such as fungi, life as we know it would cease. (See Concept 55.4 for a detailed discussion of chemical cycles.)

Prokaryotes also convert some molecules to forms that can be taken up by other organisms. Cyanobacteria and other autotrophic prokaryotes use CO_2 to make organic compounds such as sugars, which are then passed up through food chains. Cyanobacteria also produce atmospheric O_2, and a variety of prokaryotes fix atmospheric nitrogen (N_2) into forms that other organisms can use to make the building blocks of proteins and nucleic acids. Under some conditions, prokaryotes can increase the availability of nutrients that plants require for growth, such as nitrogen, phosphorus, and potassium **(Figure 27.18)**. Prokaryotes can also *decrease* the availability of key plant nutrients; this occurs when prokaryotes "immobilize" nutrients by using them to synthesize molecules that remain within their cells. Thus, prokaryotes can have complex effects on soil nutrient concentrations. In marine environments, an archaean from the clade Crenarchaeota can perform nitrification, a key step in the nitrogen cycle (see Figure 55.14). Crenarchaeotes dominate the oceans by numbers, comprising an estimated 10^{28} cells. The sheer abundance of these organisms suggests that they may have a large impact on the global nitrogen cycle.

Figure 27.18 Impact of bacteria on soil nutrient availability. Pine seedlings grown in sterile soils to which one of three strains of the bacterium *Burkholderia glathei* had been added absorbed more potassium (K^+) than did seedlings grown in soil without any bacteria. Other results (not shown) demonstrated that strain 3 increased the amount of K^+ released from mineral crystals to the soil.



WHAT IF? ➤ *Estimate the average uptake of K^+ for seedlings in soils with bacteria. What would you expect this average to be if bacteria had no effect on nutrient availability?*

Figure 27.19 Mutualism: bacterial "headlights." The glowing oval below the eye of the flashlight fish (*Photoblepharon palpebratus*) is an organ harboring bioluminescent bacteria. The fish uses the light to attract prey and to signal potential mates. The bacteria receive nutrients from the fish.

Ecological Interactions

Prokaryotes play a central role in many ecological interactions. Consider **symbiosis** (from a Greek word meaning "living together"), an ecological relationship in which two species live in close contact with each other. Prokaryotes often form symbiotic associations with much larger organisms. In general, the larger organism in a symbiotic relationship is known as the **host**, and the smaller is known as the **symbiont**. There are many cases in which a prokaryote and its host participate in **mutualism**, an ecological interaction between two species in which both benefit **(Figure 27.19)**. Other interactions take the form of **commensalism**, an ecological relationship in which one species benefits while the other is not harmed or helped in any significant way. For example, more than 150 bacterial species live on the outer surface of your body, covering portions of your skin with up to 10 million cells per square centimeter. Some of these species are commensalists: You provide them with food, such as the oils that exude from your pores, and a place to live, while they neither harm nor benefit you. Finally, some prokaryotes engage in **parasitism**, an ecological relationship in which a **parasite** eats the cell contents, tissues, or body fluids of its host. As a group, parasites harm but usually do not kill their host, at least not immediately (unlike a predator). Parasites that cause disease are known as **pathogens**, many of which are prokaryotic. (We'll discuss mutualism, commensalism, and parasitism in greater detail in Concept 54.1.)

The very existence of an ecosystem can depend on prokaryotes. For example, consider the diverse ecological communities found at hydrothermal vents. These communities are densely populated by many different kinds of animals, including worms, clams, crabs, and fishes. But since sunlight does not penetrate to the deep ocean floor, the community does not include photosynthetic organisms. Instead, the energy that supports the community is derived from the metabolic activities of chemoautotrophic bacteria. These bacteria harvest chemical energy from compounds such as hydrogen sulfide (H_2S) that are released from the vent. An active hydrothermal vent may support hundreds of eukaryotic species, but when the vent stops releasing chemicals, the chemoautotrophic bacteria cannot survive. As a result, the entire vent community collapses.

CONCEPT CHECK 27.5

1. Explain how prokaryotes, though small, can be considered giants in their collective impact on Earth and its life.
2. **MAKE CONNECTIONS ➤** Review Figure 11.6. Then summarize the main steps by which cyanobacteria produce O_2 and use CO_2 to make organic compounds.

For suggested answers, see Appendix A.

CONCEPT 27.6

Prokaryotes have both beneficial and harmful impacts on humans

Although the best-known prokaryotes tend to be the bacteria that cause human illness, these pathogens represent only a small fraction of prokaryotic species. Many other prokaryotes have positive interactions with people, and some play essential roles in agriculture and industry.

Mutualistic Bacteria

As is true for many other eukaryotes, human well-being can depend on mutualistic prokaryotes. For example, our intestines are home to an estimated 500–1,000 species of bacteria; their cells outnumber all human cells in the body by a factor of ten. Different species live in different portions of the intestines, and they vary in their ability to process different foods. Many of these species are mutualists, digesting food that our own intestines cannot break down. The genome of one of these gut mutualists, *Bacteroides thetaiotaomicron*, includes a large array of genes involved in synthesizing carbohydrates, vitamins, and other nutrients needed by humans. Signals from the bacterium activate human genes that build the network of intestinal blood vessels necessary to absorb nutrient molecules. Other signals induce human cells to produce antimicrobial compounds to which *B. thetaiotaomicron* is not susceptible. This action may reduce the population sizes of other, competing species, thus potentially benefiting both *B. thetaiotaomicron* and its human host.

Pathogenic Bacteria

All the pathogenic prokaryotes known to date are bacteria, and they deserve their negative reputation. Bacteria cause about half of all human diseases. For example, more than 1 million people die each year of the lung disease tuberculosis, caused by *Mycobacterium tuberculosis*. And another

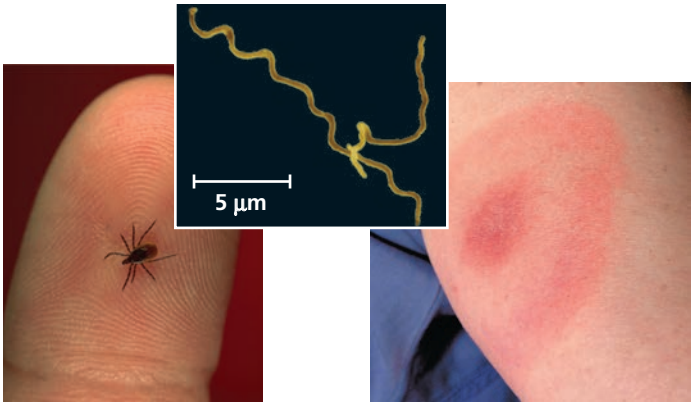


Figure 27.20 Lyme disease. Ticks in the genus *Ixodes* spread the disease by transmitting the spirochete *Borrelia burgdorferi* (colorized SEM). A rash may develop at the site of the tick's bite; the rash may be large and ring-shaped (as shown) or much less distinctive.

2 million people die each year from diarrheal diseases caused by various bacteria.

Some bacterial diseases are transmitted by other species, such as fleas or ticks. In the United States, the most widespread pest-carried disease is Lyme disease, which infects 15,000 to 20,000 people each year **(Figure 27.20)**. Caused by a bacterium carried by ticks that live on deer and field mice, Lyme disease can result in debilitating arthritis, heart disease, nervous disorders, and death if untreated.

Pathogenic prokaryotes usually cause illness by producing poisons, which are classified as exotoxins or endotoxins. **Exotoxins** are proteins secreted by certain bacteria and other organisms. Cholera, a dangerous diarrheal disease, is caused by an exotoxin secreted by the proteobacterium *Vibrio cholerae*. The exotoxin stimulates intestinal cells to release chloride ions into the gut, and water follows by osmosis. In another example, the potentially fatal disease botulism is caused by botulinum toxin, an exotoxin secreted by the gram-positive bacterium *Clostridium botulinum* as it ferments various foods, including improperly canned meat, seafood, and vegetables. Like other exotoxins, the botulinum toxin can produce disease even if the bacteria that manufacture it are no longer present when the food is eaten. Another species in the same genus, *C. difficile*, produces exotoxins that cause severe diarrhea, resulting in more than 12,000 deaths per year in the United States alone.

Endotoxins are lipopolysaccharide components of the outer membrane of gram-negative bacteria. In contrast to exotoxins, endotoxins are released only when the bacteria die and their cell walls break down. Endotoxin-producing bacteria include species in the genus *Salmonella*, such as *Salmonella typhi*, which causes typhoid fever. You might have heard of food poisoning caused by other *Salmonella* species that can be found in poultry and some fruits and vegetables.

Since the 19th century, improved sanitation systems in the industrialized world have greatly reduced the threat of pathogenic bacteria. Antibiotics have saved a great many lives and reduced the incidence of disease. However, resistance to antibiotics is currently evolving in many bacterial strains. As you read earlier, the rapid reproduction of bacteria enables cells carrying resistance genes to quickly give rise to large populations as a result of natural selection, and these genes can also spread to other species by horizontal gene transfer.

Horizontal gene transfer can also spread genes associated with virulence, turning normally harmless bacteria into potent pathogens. *E. coli*, for instance, is ordinarily a harmless symbiont in the human intestines, but pathogenic strains that cause bloody diarrhea have emerged. One of the most dangerous strains, O157:H7, is a global threat; in the United States alone, there are 75,000 cases of O157:H7 infection per year, often from contaminated beef or produce. Scientists have sequenced the genome of O157:H7 and compared it with the genome of a harmless strain of *E. coli* called K-12. They discovered that 1,387 out of the 5,416 genes in O157:H7 have no counterpart in K-12. Many of these 1,387 genes are found in chromosomal regions that include phage DNA. This suggests that at least some of the 1,387 genes were incorporated into the genome of O157:H7 through phage-mediated horizontal gene transfer (transduction). Some of the genes found only in O157:H7 are associated with virulence, including genes that code for adhesive fimbriae that enable O157:H7 to attach itself to the intestinal wall and extract nutrients.

Prokaryotes in Research and Technology

On a positive note, we reap many benefits from the metabolic capabilities of both bacteria and archaea. For example, people have long used bacteria to convert milk to cheese and yogurt. Bacteria are also used in the production of beer and wine, pepperoni, fermented cabbage (sauerkraut), and soy sauce. In recent decades, our greater understanding of prokaryotes has led to an explosion of new applications in biotechnology. Examples include the use of *E. coli* in gene cloning (see Figure 19.4) and the use of DNA polymerase from *Pyrococcus furiosus* in the PCR technique (see Figure 19.8). Through genetic engineering, we can modify bacteria to produce vitamins, antibiotics, hormones, and other products (see Concept 19.1). In addition, naturally occurring soil bacteria have potential as sources of new antibiotics, as you can explore in the **Scientific Skills Exercise**.

Recently, the prokaryotic CRISPR-Cas system, which helps bacteria and archaea defend against attack by viruses (see Figure 26.7), has been developed into a powerful new tool for altering genes in virtually any organism. The genomes of many prokaryotes contain short DNA repeats, called CRISPRs, that interact with proteins known as the Cas (CRISPR-associated) proteins. Cas proteins, acting together with "guide RNA" made from the CRISPR region, can cut any DNA sequence to which they are directed. Scientists have been able to exploit this system by introducing a Cas protein (Cas9) to guide RNA into cells whose DNA they want to alter (see Figure 19.14). Among

SCIENTIFIC SKILLS EXERCISE

Calculating and Interpreting Means and Standard Errors

Can Antibiotics Obtained from Soil Bacteria Help Fight Drug-Resistant Bacteria? Soil bacteria synthesize antibiotics, which they use against species that attack or compete with them. To date, these species have been inaccessible as sources for new medicines because 99% of soil bacteria cannot be grown using standard laboratory techniques. To address this problem, researchers developed a method in which soil bacteria grow in a simulated version of their natural environment; this led to the discovery of a new antibiotic, teixobactin. In this exercise, you'll calculate means and standard errors from an experiment that tested teixobactin's effectiveness against MRSA (methicillin-resistant *Staphylococcus aureus*; see Figure 21.14).

How the Experiment Was Done Researchers drilled tiny holes into a small plastic chip and filled the holes with a dilute aqueous solution containing soil bacteria and agar. The dilution had been calibrated so that only one bacterium was likely to grow in each hole. After the agar solidified, the chip was then placed in a container containing the original soil; nutrients and other essential materials from the soil diffused into the agar, allowing the bacteria to grow.

After isolating teixobactin from a soil bacterium, researchers performed the following experiment: mice infected with MRSA were given low (1 mg/kg) or high (5 mg/kg) doses of teixobactin or vancomycin, an existing antibiotic; in the control, mice infected with MRSA were not given an antibiotic. After 26 hours, researchers sampled infected mice and estimated the number of *S. aureus* colonies in each sample. Results were reported on a log scale; note that a decrease of 1.0 on this scale reflects a 10-fold decrease in MRSA abundance.

Data from the Experiment

Treatment	Dose (mg/kg)	Log of Number of Colonies	Mean ($\bar{x}$)
Control	—	9.0, 9.5, 9.0, 8.9	
Vancomycin	1.0	8.5, 8.4, 8.2	
	5.0	5.3, 5.9, 4.7	
Teixobactin	1.0	8.5, 6.0, 8.4, 6.0	
	5.0	3.8, 4.9, 5.2, 4.9	

➤ Plastic chip used to grow soil bacteria

Data from L. Ling et al. A new antibiotic kills pathogens without detectable resistance, *Nature* 517:455–459 (2015).

INTERPRET THE DATA

1. The mean ($\bar{x}$) of a variable is the sum of the data values divided by the number of observations (n):

$$\bar{x} = \frac{\sum x_i}{n}$$

In this formula, x_i is the value of the ith observation of the variable; the $\sum$ symbol indicates that the n values of x are to be added together. Calculate the mean for each treatment.

2. Use your results from question 1 to evaluate the effectiveness of vancomycin and teixobactin.
3. The variation found in a set of data can be estimated by the standard deviation, s:

$$s = \sqrt{\frac{1}{n-1}\sum (x_i - \bar{x})^2}$$

Calculate the standard deviation for each treatment.

4. The standard error (SE), which indicates how greatly the mean would likely vary if the experiment was repeated, is calculated as:

$$\text{SE} = \frac{s}{\sqrt{n}}$$

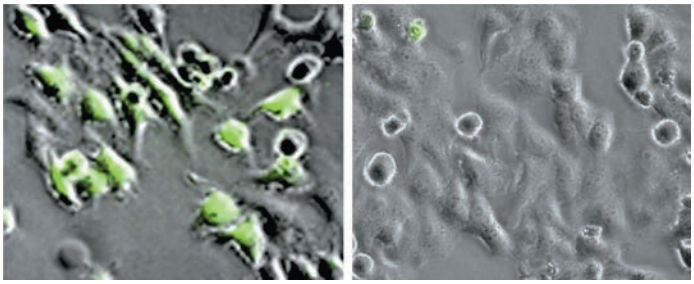
As a rough rule of thumb, if an experiment were to be repeated, the new mean typically would lie within two standard errors of the original mean (that is, within the range $\bar{x} \pm 2\text{SE}$). Calculate $\bar{x} \pm 2\text{SE}$ for each treatment, determine whether these ranges overlap, and interpret your results.

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

other applications, this **CRISPR-Cas9 system** has already opened new lines of research on HIV, the virus that causes AIDS **(Figure 27.21)**. While the CRISPR-Cas9 system can potentially be used in many different ways, care must be taken to guard against the unintended consequences that could arise when applying such a new and powerful technology.

Another valuable application of bacteria is to reduce our use of petroleum. Consider the plastics industry. Globally, each year about 350 billion pounds of plastic are produced from petroleum and used to make toys, storage containers, soft drink bottles, and many other items. These products degrade slowly, creating environmental problems. Bacteria produce natural plastics **(Figure 27.22)**. For example, some bacteria synthesize a type of polymer known as PHA (polyhydroxyalkanoate), which they use to store chemical energy. The PHA can be extracted, formed into pellets, and used to make durable, yet biodegradable, plastics.

➤ Figure 27.21 CRISPR: Opening new avenues of research for treating HIV infection. (a) In laboratory experiments, untreated (control) human cells were susceptible to infection by HIV, the virus that causes AIDS. **(b)** In contrast, cells treated with a CRISPR-Cas9 system that targets HIV were resistant to viral infection. The CRISPR-Cas9 system was also able to remove HIV proviruses that had become incorporated into the DNA of human cells.



(a) Control cells. The green color indicates infection by HIV.

(b) Experimental cells. These cells were treated with a CRISPR-Cas9 system that targets HIV.

Researchers are also seeking to reduce the use of petroleum and other fossil fuels by engineering bacteria that can produce ethanol from various forms of biomass, including agricultural waste, switchgrass, and corn.

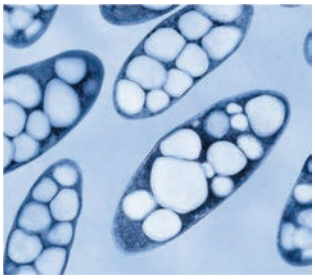


Figure 27.22 Bacteria synthesizing and storing PHA, a component of biodegradable plastics.

Another way to harness prokaryotes is **bioremediation**, the use of organisms to remove pollutants from soil, air, or water. For example, anaerobic bacteria and archaea decompose the organic matter in sewage, converting it to material that can be used as landfill or fertilizer after chemical sterilization. Other bioremediation applications include cleaning up oil spills **(Figure 27.23)** and precipitating radioactive material (such as uranium) out of groundwater.

The usefulness of prokaryotes largely derives from their diverse forms of nutrition and metabolism. All this metabolic versatility evolved prior to the appearance of the structural novelties that heralded the evolution of eukaryotic organisms, discussed in the rest of this unit.

Figure 27.23 Bioremediation of an oil spill. Spraying fertilizer stimulates the growth of native bacteria that metabolize oil, increasing the speed of the breakdown process up to fivefold.

CONCEPT CHECK 27.6

1. Identify at least two ways that prokaryotes have affected you positively today.
2. A pathogenic bacterium's toxin causes symptoms that increase the bacterium's chance of spreading from host to host. Does this information indicate whether the poison is an exotoxin or endotoxin? Explain.
3. **WHAT IF?** How might a sudden and dramatic change in your diet affect the diversity of prokaryotic species that live in your digestive tract?

For suggested answers, see Appendix A.

27 Chapter Review

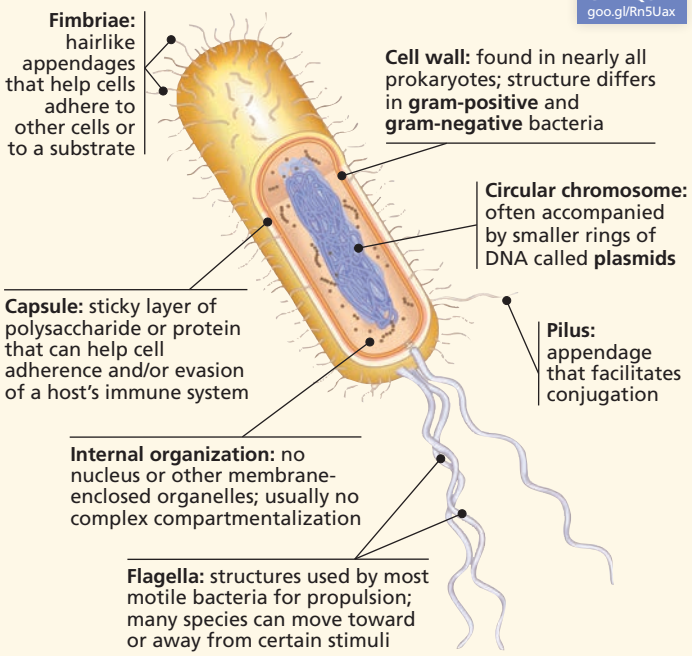
Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 27.1

Structural and functional adaptations contribute to prokaryotic success *(pp. 626–630)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax



- Many prokaryotic species can reproduce quickly by binary fission, leading to the formation of extremely large populations.

? *Describe features of prokaryotes that enable them to thrive in a wide range of different environments.*

CONCEPT 27.2

Rapid reproduction, mutation, and genetic recombination promote genetic diversity in prokaryotes *(pp. 630–633)*

- Because prokaryotes can often proliferate rapidly, mutations can quickly increase a population's genetic variation. As a result, prokaryotic populations often can evolve in short periods of time in response to changing conditions.
- Genetic diversity in prokaryotes also can arise by recombination of the DNA from two different cells (via transformation, transduction, or conjugation). By transferring advantageous alleles, such as ones for antibiotic resistance, recombination can promote adaptive evolution in prokaryotic populations.

? *Mutations are rare and prokaryotes reproduce asexually, yet their populations can have high genetic diversity. Explain how this can occur.*

CONCEPT 27.3

Diverse nutritional and metabolic adaptations have evolved in prokaryotes *(pp. 633–634)*

- Nutritional diversity is much greater in prokaryotes than in eukaryotes and includes all four modes of nutrition: photoautotrophy, chemoautotrophy, photoheterotrophy, and chemoheterotrophy.

- Among prokaryotes, **obligate aerobes** require O_2, **obligate anaerobes** are poisoned by O_2, and **facultative anaerobes** can survive with or without O_2.
- Unlike eukaryotes, prokaryotes can metabolize nitrogen in many different forms. Some can convert atmospheric nitrogen to ammonia, a process called **nitrogen fixation**.
- Prokaryotic cells and even species may cooperate metabolically. Metabolic cooperation also occurs in surface-coating **biofilms** that include different species.

Describe the range of prokaryotic metabolic adaptations.

CONCEPT 27.4

Prokaryotes have radiated into a diverse set of lineages *(pp. 635–639)*

- Molecular systematics is helping biologists classify prokaryotes and identify new clades.
- Diverse nutritional types are scattered among the major groups of bacteria. The two largest groups are the proteobacteria and gram-positive bacteria.
- Some archaea, such as extreme thermophiles and extreme halophiles, live in extreme environments. Other archaea live in moderate environments such as soils and lakes.

How have molecular data informed prokaryotic phylogeny?

CONCEPT 27.5

Prokaryotes play crucial roles in the biosphere *(pp. 639–640)*

- Decomposition by heterotrophic prokaryotes and the synthetic activities of autotrophic and nitrogen-fixing prokaryotes contribute to the recycling of elements in ecosystems.
- Many prokaryotes have a symbiotic relationship with a host; the relationships between prokaryotes and their hosts range from mutualism to commensalism to parasitism.

In what ways are prokaryotes key to the survival of many species?

CONCEPT 27.6

Prokaryotes have both beneficial and harmful impacts on humans *(pp. 640–643)*

- People depend on mutualistic prokaryotes, including hundreds of species that live in our intestines and help digest food.
- Pathogenic bacteria typically cause disease by releasing **exotoxins** or **endotoxins**. Horizontal gene transfer can spread genes associated with virulence to harmless species or strains.
- Prokaryotes can be used in **bioremediation** and production of plastics, vitamins, antibiotics, and other products.

Describe beneficial and harmful impacts of prokaryotes on humans.

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–6 can be found in the Study Area in MasteringBiology.

PRACTICE TEST
goo.gl/iAsVgL

7. **EVOLUTION CONNECTION** In patients with nonresistant strains of the tuberculosis bacterium, antibiotics can relieve symptoms in a few weeks. However, it takes much longer to halt the infection, and patients may discontinue treatment while bacteria are still present. Explain how this could result in the evolution of drug-resistant pathogens.

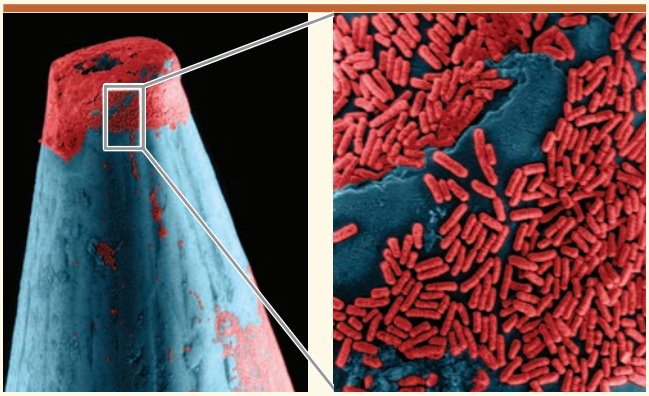
8. **SCIENTIFIC INQUIRY • INTERPRET THE DATA** The nitrogen-fixing bacterium *Rhizobium* infects the roots of some plant species, forming a mutualism in which the bacterium provides nitrogen, and the plant provides carbohydrates. Scientists measured the 12-week growth of one such plant species (*Acacia irrorata*) when infected by six different *Rhizobium* strains. (a) Graph the data. (b) Interpret your graph.

Rhizobium strain	1	2	3	4	5	6
Plant mass (g)	0.91	0.06	1.56	1.72	0.14	1.03

Data from J. J. Burdon et al., Variation in the effectiveness of symbiotic associations between native rhizobia and temperate Australian *Acacia*: within species interactions, *Journal of Applied Ecology* 36:398–408 (1999).
Note: Without *Rhizobium*, after 12 weeks, *Acacia* plants have a mass of about 0.1 g.

9. **WRITE ABOUT A THEME: ENERGY** In a short essay (about 100–150 words), discuss how prokaryotes and other members of hydrothermal vent communities transfer and transform energy.

10. **SYNTHESIZE YOUR KNOWLEDGE**



Explain how the small size and rapid reproduction rate of bacteria (such as the population shown here on the tip of a pin) contribute to their large population sizes and high genetic variation.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

The Origin and Evolution of Eukaryotes

28

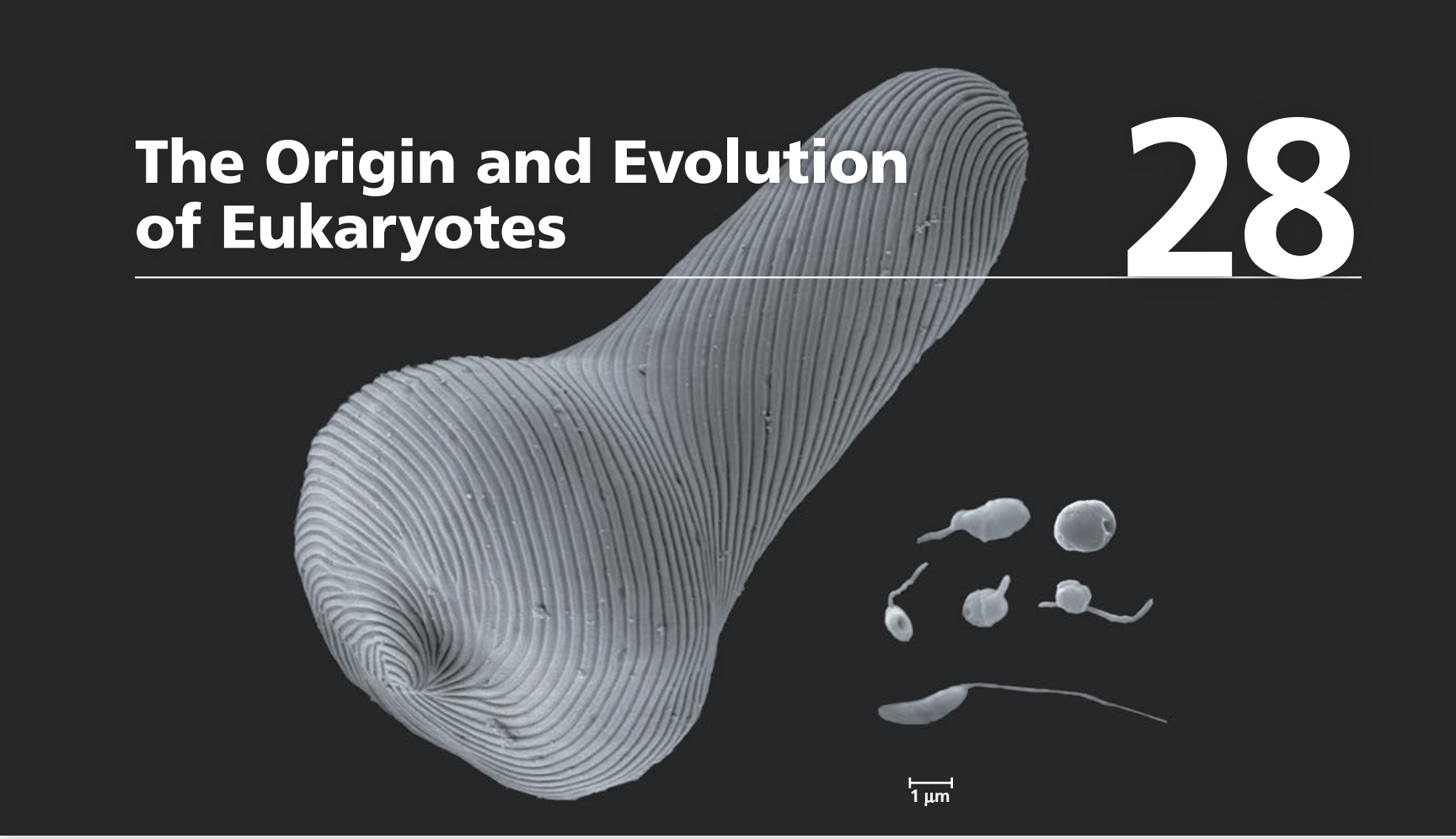


Figure 28.1 Which of these organisms are prokaryotes and which are eukaryotes?

KEY CONCEPTS

Trumpet-shaped protists (*Stentor coeruleus*)

Living Small

Knowing that most prokaryotes are extremely small organisms, you might assume that **Figure 28.1** depicts six prokaryotes and one much larger eukaryote. But in fact, the only prokaryote is the organism immediately above the scale bar. The other six organisms are members of diverse, mostly unicellular groups of eukaryotes informally known as **protists**. These very small eukaryotes have intrigued biologists for more than 300 years, ever since the Dutch scientist Antoni van Leeuwenhoek first laid eyes on them under a light microscope. Some protists change their forms as they creep along using blob-like appendages, while others resemble tiny trumpets or miniature jewelry. Recalling his observations, van Leeuwenhoek wrote, "No more pleasant sight has met my eye than this, of so many thousands of living creatures in one small drop of water."

The protists that fascinated van Leeuwenhoek continue to surprise us today. Metagenomic studies have revealed a treasure trove of previously unknown protists within the world of microscopic life. Many of these newly discovered organisms are just 0.5–2 µm in diameter—as small as many prokaryotes. Genetic and morphological studies have also shown that some protists are more closely related to plants, fungi, or animals than they are to other protists. As a result, the kingdom in which all protists once were classified, Protista, has been abandoned, and various protist lineages are now recognized as major groups in their own right. Most biologists still

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

use the term *protist*, but only as a convenient way to refer to eukaryotes that are not plants, animals, or fungi.

In this chapter, you will become acquainted with some of the most significant groups of protists. You will learn about their structural and biochemical adaptations as well as their enormous impact on ecosystems, agriculture, industry, and human health.

HHMI Video: Seeing the Invisible: Van Leeuwenhoek's First Glimpses of the Microbial World

CONCEPT 28.1

Most eukaryotes are single-celled organisms

Protists, along with plants, animals, and fungi, are classified as eukaryotes; they are in domain Eukarya, one of the three domains of life. Unlike the cells of prokaryotes, eukaryotic cells have a nucleus and other membrane-enclosed organelles, such as mitochondria and the Golgi apparatus. Such organelles provide specific locations where particular cellular functions are accomplished, making the structure and organization of eukaryotic cells more complex than those of prokaryotic cells.

Eukaryotic cells also have a well-developed cytoskeleton that extends throughout the cell (see Figure 7.20). The cytoskeleton provides the structural support that enables eukaryotic cells to have asymmetric (irregular) forms, as well as to change in shape as they feed, move, or grow. In contrast, prokaryotic cells lack a well-developed cytoskeleton, thus limiting the extent to which they can maintain asymmetric forms or change shape over time.

We'll survey the diversity of eukaryotes throughout the rest of this unit, beginning in this chapter with the protists. As you explore this material, bear in mind that

- the organisms in most eukaryotic lineages are protists, and
- most protists are unicellular.

Thus, life differs greatly from how most of us commonly think of it. The large, multicellular organisms that we know best (plants, animals, and fungi) are the tips of just a few branches on the great tree of life (see Figure 22.21).

Structural and Functional Diversity in Protists

Given that they are classified in a number of different groups, it isn't surprising that few general characteristics of protists can be cited without exceptions. In fact, protists exhibit more structural and functional diversity than the eukaryotes with which we are most familiar—plants, animals, and fungi.

For example, most protists are unicellular, although there are some colonial and multicellular species. Single-celled protists are justifiably considered the simplest eukaryotes, but at the cellular level, many protists are very complex—the most elaborate of all cells. In multicellular organisms, essential biological functions are carried out by organs. Unicellular protists carry out the same essential functions, but they do so using subcellular organelles, not multicellular organs. The organelles that protists use are mostly those discussed in Figure 7.8, including the nucleus, endoplasmic reticulum, Golgi apparatus, and lysosomes. Certain protists also rely on organelles not found in most other eukaryotic cells, such as contractile vacuoles that pump excess water from the protistan cell (see Figure 8.13).

Protists are also very diverse in their nutrition. Some protists are photoautotrophs and contain chloroplasts. Some are heterotrophs, absorbing organic molecules or ingesting larger food particles. Still other protists, called **mixotrophs**, combine photosynthesis and heterotrophic nutrition. Photoautotrophy, heterotrophy, and mixotrophy have all arisen independently in many different protist lineages.

Reproduction and life cycles also are highly varied among protists. Some protists are only known to reproduce asexually; others can also reproduce sexually or at least employ the sexual processes of meiosis and fertilization. All three basic types of sexual life cycles (see Figure 13.6) are represented among protists, along with some variations that do not quite fit any of these types. We will examine the life cycles of several protist groups later in this chapter.

Four Supergroups of Eukaryotes

Our understanding of the evolutionary history of eukaryotic diversity has been in flux in recent years. Not only has kingdom Protista been abandoned, but other hypotheses have been discarded as well. For example, many biologists once thought that the first lineage to have diverged from all other eukaryotes was the *amitochondriate protists*, organisms without conventional mitochondria and with fewer membrane-enclosed organelles than other protist groups. But recent structural and DNA data have undermined this hypothesis. Many of the so-called amitochondriate protists have been shown to have mitochondria—though reduced ones—and some of these organisms are now classified in distantly related groups.

The ongoing changes in our understanding of the phylogeny of protists pose challenges to students and instructors alike. Hypotheses about these relationships are a focus of scientific activity, changing rapidly as new data cause previous ideas to be modified or discarded. We'll focus here on one current hypothesis: the four supergroups of eukaryotes shown in **Figure 28.2**. Because the root of the eukaryotic tree is not known, all four supergroups are shown as diverging simultaneously from a common ancestor. We know that this is not correct, but we do not know which supergroup was the first to diverge from the others. In addition, while some of the groups in Figure 28.2 are well supported by morphological and DNA data, others are more controversial. As you read this chapter, it may be helpful to focus less on the specific names

of groups of organisms and more on why the organisms are important and how ongoing research is elucidating their evolutionary relationships.

Endosymbiosis in Eukaryotic Evolution

What gave rise to the enormous diversity of protists that exist today? There is abundant evidence that much of protistan diversity has its origins in **endosymbiosis**, a relationship between two species in which one organism lives inside the cell or cells of another organism (the host). In particular, as we discussed in Concept 25.3, structural, biochemical, and DNA sequence data indicate that mitochondria and plastids are derived from prokaryotes that were engulfed by the ancestors of early eukaryotic cells. The evidence also suggests that mitochondria evolved before plastids. Thus, a defining moment in the origin of eukaryotes occurred when a host cell engulfed a bacterium that would later become an organelle found in all eukaryotes—the mitochondrion.

To determine which prokaryotic lineage gave rise to mitochondria, researchers have compared the DNA sequences of mitochondrial genes (mtDNA) to those found in major clades of bacteria and archaea. In the **Scientific Skills Exercise**, you will interpret one such set of DNA sequence comparisons.

SCIENTIFIC SKILLS EXERCISE

Interpreting Comparisons of Genetic Sequences

Which Prokaryotes Are Most Closely Related to Mitochondria? Early eukaryotes acquired mitochondria by endosymbiosis: A host cell engulfed an aerobic prokaryote that persisted within the cytoplasm to the mutual benefit of both cells. In studying which living prokaryotes might be most closely related to mitochondria, researchers compared ribosomal RNA (rRNA) sequences. Ribosomes perform critical cell functions. Hence, rRNA sequences are under strong selection and change slowly over time, making them suitable for comparing even distantly related species. In this exercise, you will interpret some of the research data to draw conclusions about the phylogeny of mitochondria.

How the Research Was Done Researchers isolated and cloned nucleotide sequences from the gene that codes for the small-subunit rRNA molecule for wheat (a eukaryote) and five bacterial species:

- Wheat, used as the source of mitochondrial rRNA genes
- *Agrobacterium tumefaciens*, an alpha proteobacterium that lives within plant tissue and produces tumors in the host
- *Comamonas testosteroni*, a beta proteobacterium
- *Escherichia coli*, a well-studied gamma proteobacterium that inhabits human intestines
- *Mycoplasma capricolum*, a gram-positive mycoplasma, which is the only group of bacteria lacking cell walls
- *Anacystis nidulans*, a cyanobacterium

Data from the Research Cloned rRNA gene sequences for the six organisms were aligned and compared. The data table below, called a *comparison matrix*, summarizes the comparison of 617 nucleotide positions from the gene sequences. Each value in the table is the percentage of the 617 nucleotide positions for which the pair of organisms have the same base. Any positions that were identical across the rRNA genes of all six organisms were omitted from this comparison matrix.

Wheat, used as the source of mitochondrial RNA

INTERPRET THE DATA

1. First, make sure you understand how to read the comparison matrix. Find the cell that represents the comparison of *C. testosteroni* and *E. coli*. What value is given in this cell? What does that value signify about the comparable rRNA gene sequences in those two organisms? Explain why some cells have a dash rather than a value. Why are some cells shaded gray, with no value?
2. Why did the researchers choose one plant mitochondrion and five bacterial species to include in the comparison matrix?
3. Which bacterium has an rRNA gene that is most similar to that of the wheat mitochondrion? What is the significance of this similarity?

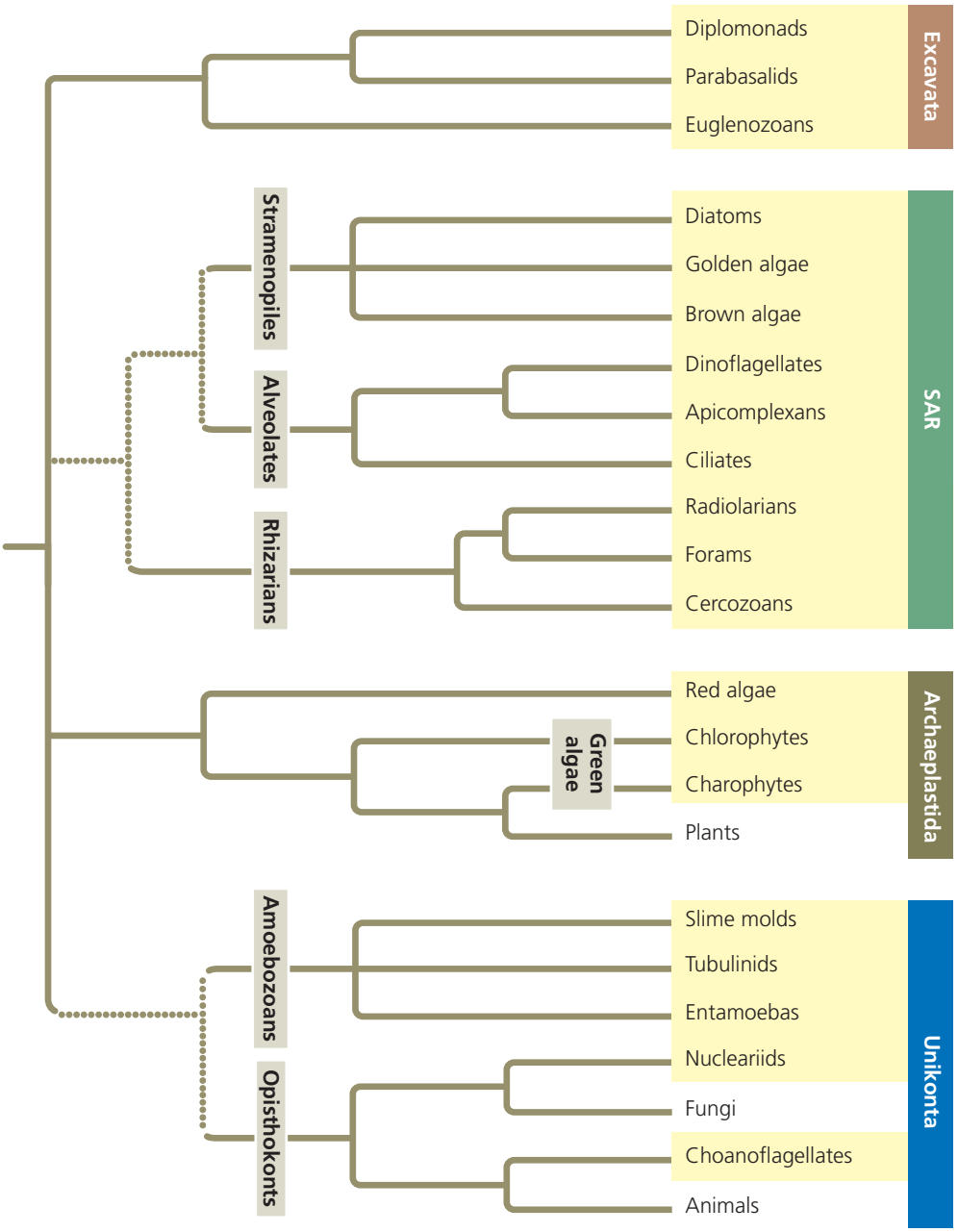
Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

	Wheat mitochondrion	*A. tumefaciens*	*C. testosteroni*	*E. coli*	*M. capricolum*	*A. nidulans*
Wheat mitochondrion	–	48	38	35	34	34
A. tumefaciens		–	55	57	52	53
C. testosteroni			–	61	52	52
E. coli				–	48	52
M. capricolum					–	50
A. nidulans						–

Data from D. Yang et al., Mitochondrial origins, *Proceedings of the National Academy of Sciences USA* 82:4443–4447 (1985).

Figure 28.2 Exploring Protistan Diversity

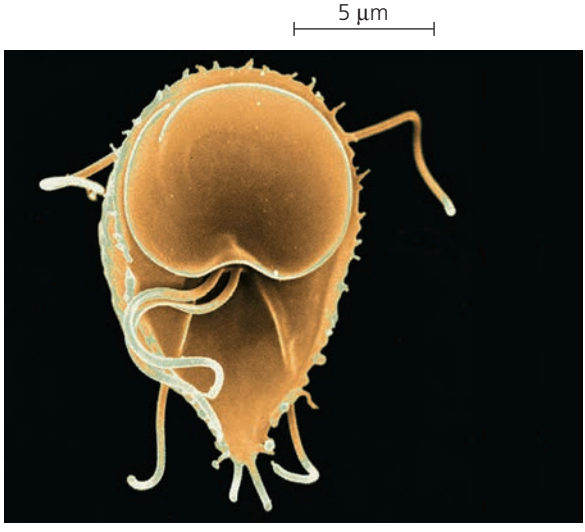
The tree below represents a phylogenetic hypothesis for the relationships among eukaryotes on Earth today. The eukaryotic groups at the branch tips are related in larger "supergroups," labeled vertically at the far right of the tree. Groups that were formerly classified in the kingdom Protista are highlighted in yellow. Dotted lines indicate evolutionary relationships that are uncertain and proposed clades that are under active debate. For clarity, this tree only includes representative clades from each supergroup. In addition, the recent discoveries of many new groups of eukaryotes indicate that eukaryotic diversity is actually much greater than shown here.



DRAW IT ➤ *Draw a simplified version of this phylogenetic tree that shows only the four supergroups of eukaryotes. Now sketch how the tree would look if the unikonts were the sister group to all other eukaryotes.*

Excavata

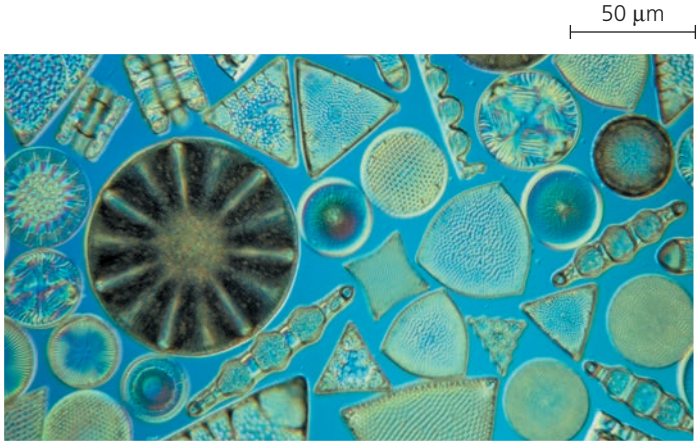
Some members of this supergroup have an "excavated" groove on one side of the cell body. Two major clades (the parabasalids and diplomonads) have highly reduced mitochondria; members of a third clade (the euglenozoans) have flagella that differ in structure from those of other organisms. Excavates include parasites such as *Giardia*, as well as many predatory and photosynthetic species.



***Giardia intestinalis*, a diplomonad parasite.** This diplomonad (colorized SEM), which lacks the characteristic surface groove of the Excavata, inhabits the intestines of mammals. It can infect people when they drink water contaminated with feces containing *Giardia* cysts. Drinking such water—even from a seemingly pristine stream—can cause severe diarrhea. Boiling the water kills the parasite.

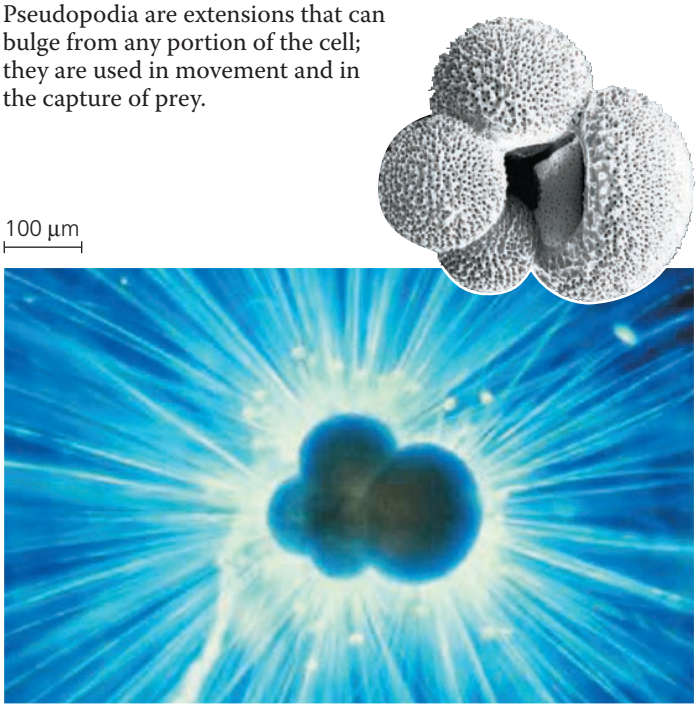
SAR

This supergroup contains (and is named after) three large and very diverse clades: Stramenopila, Alveolata, and Rhizaria. Stramenopiles include some of the most important photosynthetic organisms on Earth, such as the diatoms shown here. Alveolates also include photosynthetic species, as well as important pathogens, such as *Plasmodium*, which causes malaria. According to one current hypothesis, stramenopiles and alveolates originated by secondary endosymbiosis when a heterotrophic protist engulfed a red alga.



Diatom diversity. These beautiful single-celled protists are important photosynthetic organisms in aquatic communities (LM).

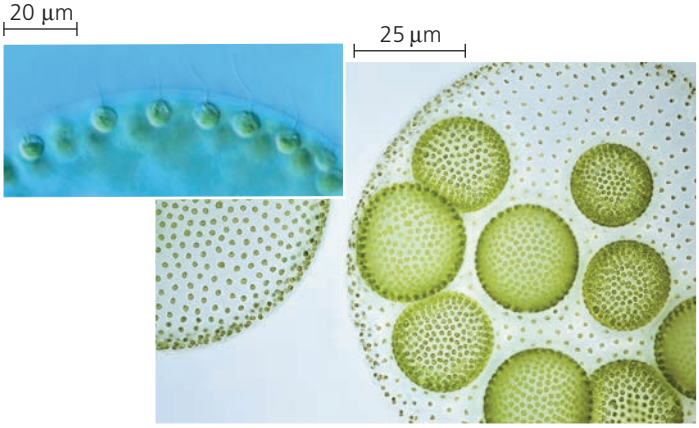
The rhizarian subgroup of SAR includes many species of amoebas, most of which have pseudopodia that are threadlike in shape. Pseudopodia are extensions that can bulge from any portion of the cell; they are used in movement and in the capture of prey.



***Globigerina*, a rhizarian in SAR.** This species is a foram, a group whose members have threadlike pseudopodia that extend through pores in the shell, or test (LM). The inset shows a foram test, which is hardened by calcium carbonate.

Archaeplastida

This supergroup of eukaryotes includes red algae and green algae, along with plants. Red algae and green algae include unicellular species, colonial species, and multicellular species (including the green alga *Volvox*). Many of the large algae known informally as "seaweeds" are multicellular red or green algae. Protists in Archaeplastida include key photosynthetic species that form the base of the food web in many aquatic communities.



***Volvox*, a multicellular freshwater green alga.** This alga has two types of differentiated cells, and so it is considered multicellular rather than colonial. It resembles a hollow ball whose wall is composed of hundreds of biflagellated cells (see inset LM) embedded in a gelatinous extracellular matrix; if isolated, these cells cannot reproduce. However, the alga also contains cells that are specialized for either sexual or asexual reproduction. The large algae shown here will eventually release the small "daughter" algae that can be seen within them (LM).

MB Video: *Volvox*

Unikonta

This supergroup of eukaryotes includes amoebas that have lobe- or tube-shaped pseudopodia, as well as animals, fungi, and non-amoeba protists that are closely related to animals or fungi. According to one current hypothesis, the unikonts were the first eukaryotic supergroup to diverge from all other eukaryotes; however, this hypothesis has yet to be widely accepted.

A unikont amoeba. This amoeba, the tubulinid *Amoeba proteus*, is using its pseudopodia to move.

MB Video: *Amoeba* Pseudopodium

Collectively, such studies indicate that mitochondria arose from an alpha proteobacterium (see Figure 27.16). Results from mtDNA sequence analyses also indicate that the mitochondria of protists, animals, fungi and plants descended from a single common ancestor, thus suggesting that mitochondria arose only once over the course of evolution. Similar analyses provide evidence that plastids descended from a single common ancestor—a cyanobacterium that was engulfed by a eukaryotic host cell.

Progress has also been made toward identifying the host cell that engulfed an alpha proteobacterium, thereby setting the stage for the origin of eukaryotes. In 2015, for example, researchers reported the discovery of a new group of archaea, the lokiarchaeotes. In phylogenomic analyses, this group was identified as the sister group of the eukaryotes and its genome was found to encode many eukaryote-specific features. Was the host cell that engulfed an alpha proteobacterium a lokiarchaeote? While this may have been the case, it is also possible that the host was closely related to the archaeans (but was not itself an archaean). Either way, current evidence indicates that the host was a relatively complex cell in which certain features of eukaryotic cells had evolved, such as a cytoskeleton that enabled it to change shape (and thereby engulf the alpha proteobacterium).

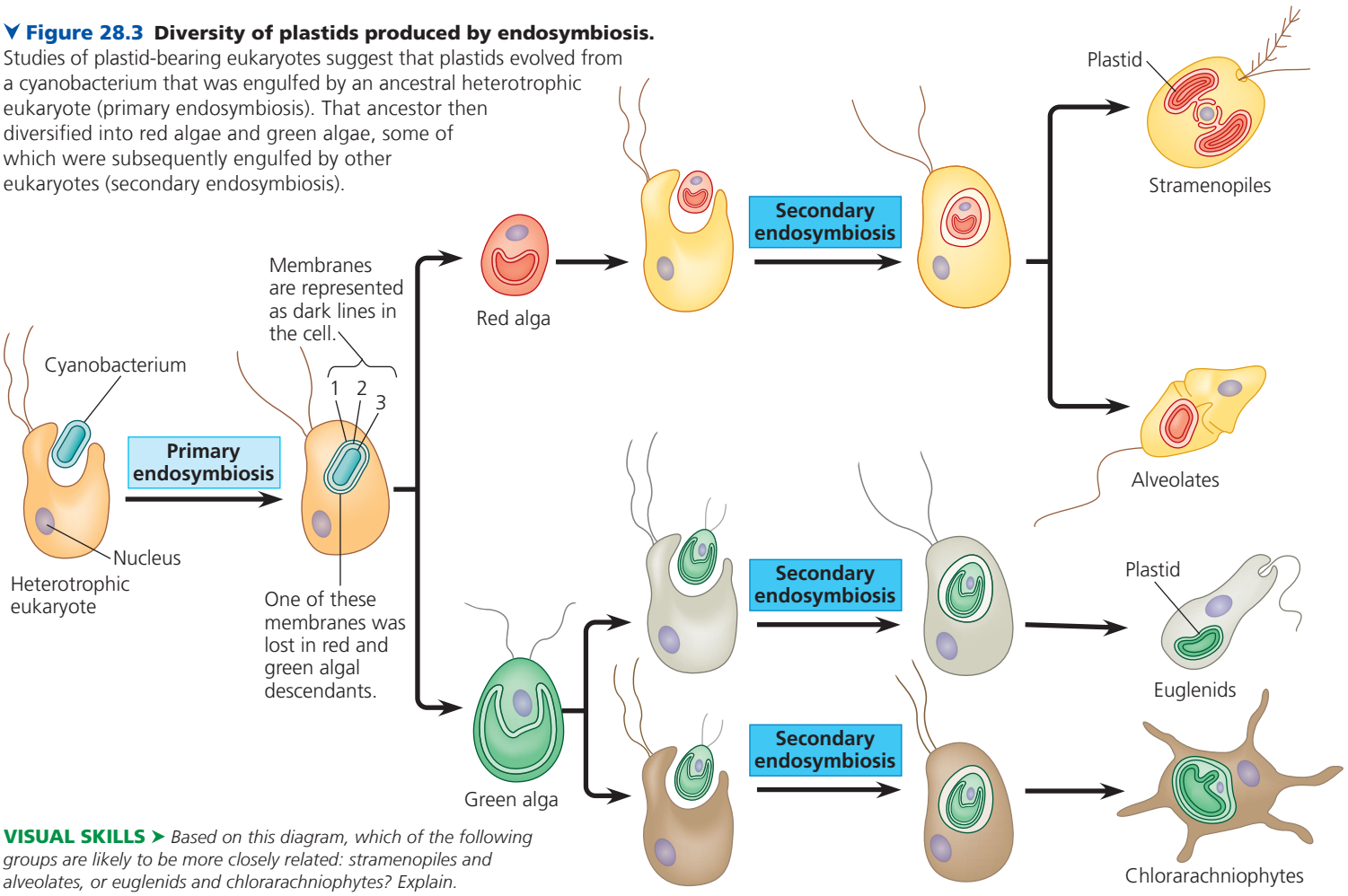
Plastid Evolution: A Closer Look

As you've seen, current evidence indicates that mitochondria are descended from a bacterium that was engulfed by a host cell that was an archaean (or a close relative of the archaeans). This event gave rise to the eukaryotes. There is also much evidence that later in eukaryotic history, a lineage of heterotrophic eukaryotes acquired an additional endosymbiont—a photosynthetic cyanobacterium—that then evolved into plastids. According to the hypothesis illustrated in **Figure 28.3**, this plastid-bearing lineage gave rise to two lineages of photosynthetic protists, or **algae**: red algae and green algae.

Let's examine some of the steps in Figure 28.3 more closely. First, recall that cyanobacteria are gram-negative and that gram-negative bacteria have two cell membranes, an inner plasma membrane and an outer membrane that is part of the cell wall (see Figure 27.3). Plastids in red algae and green algae are also surrounded by two membranes. Transport proteins in these membranes are homologous to proteins in the inner and outer membranes of cyanobacteria, providing further support for the hypothesis that plastids originated from a cyanobacterial endosymbiont.

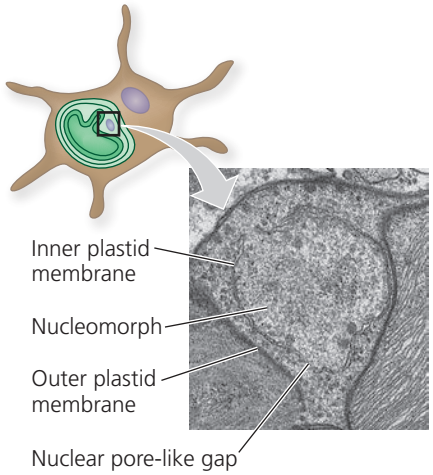
On several occasions during eukaryotic evolution, red algae and green algae underwent **secondary endosymbiosis**,

Figure 28.3 Diversity of plastids produced by endosymbiosis. Studies of plastid-bearing eukaryotes suggest that plastids evolved from a cyanobacterium that was engulfed by an ancestral heterotrophic eukaryote (primary endosymbiosis). That ancestor then diversified into red algae and green algae, some of which were subsequently engulfed by other eukaryotes (secondary endosymbiosis).



VISUAL SKILLS ➤ *Based on this diagram, which of the following groups are likely to be more closely related: stramenopiles and alveolates, or euglenids and chlorarachniophytes? Explain.*

➤ Figure 28.4 Nucleomorph within a plastid of a chlorarachniophyte.



meaning they were ingested in the food vacuoles of heterotrophic eukaryotes and became endosymbionts themselves. For example, protists known as chlorarachniophytes likely evolved when a heterotrophic eukaryote engulfed a green alga. Evidence for this process can be found within the engulfed cell, which contains a tiny vestigial nucleus, called a *nucleomorph* **(Figure 28.4)**. Genes from the nucleomorph are still transcribed, and their DNA sequences indicate that the engulfed cell was a green alga.

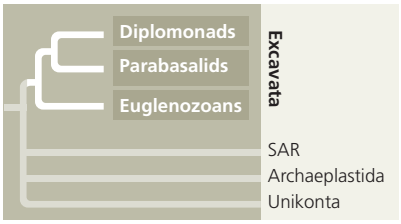
CONCEPT CHECK 28.1

1. Cite at least four examples of structural and functional diversity among protists.
2. Summarize the role of endosymbiosis in eukaryotic evolution.
3. **MAKE CONNECTIONS ➤** After studying Figure 28.3, predict how many distinct genomes are contained within the cell of a chlorarachniophyte. Explain. (See Figures 7.17 and 7.18).

For suggested answers, see Appendix A.

CONCEPT 28.2

Excavates include protists with modified mitochondria and protists with unique flagella



Now that we have examined some of the broad patterns in eukaryotic evolution, we will look more closely at the four main groups of protists shown in Figure 28.2.

We begin with **Excavata** (the excavates), a clade that was originally proposed based on morphological studies of the cytoskeleton. Some members of this diverse group also have an "excavated" feeding groove on one side of the cell body. The excavates include the diplomonads, parabasalids, and euglenozoans. Molecular data indicate that each of these three groups is monophyletic, and recent genomic studies support the monophyly of the excavate supergroup.

Diplomonads and Parabasalids

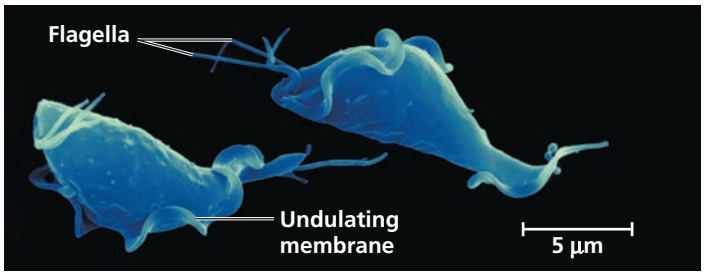
The protists in these two groups lack plastids and have highly reduced mitochondria (until recently, they were thought to lack mitochondria altogether). Most diplomonads and parabasalids are found in anaerobic environments.

Diplomonads have reduced mitochondria called *mitosomes*. These organelles lack functional electron transport chains and hence cannot use oxygen to help extract energy from carbohydrates and other organic molecules. Instead, diplomonads get the energy they need from anaerobic biochemical pathways. Many diplomonads are parasites, including the infamous *Giardia intestinalis* (see Figure 28.2), which inhabits the intestines of mammals.

Structurally, diplomonads have two equal-sized nuclei and multiple flagella. Recall that eukaryotic flagella are extensions of the cytoplasm, consisting of bundles of microtubules covered by the cell's plasma membrane (see Figure 7.24). They are quite different from prokaryotic flagella, which are filaments composed of globular proteins attached to the cell surface (see Figure 27.7).

Parabasalids also have reduced mitochondria; called *hydrogenosomes*, these organelles generate some energy anaerobically, releasing hydrogen gas as a by-product. The best-known parabasalid is *Trichomonas vaginalis*, a sexually transmitted parasite that infects some 5 million people each year. *T. vaginalis* travels along the mucus-coated lining of the human reproductive and urinary tracts by moving its flagella and by undulating part of its plasma membrane **(Figure 28.5)**. In females, if the vagina's normal acidity is disturbed, *T. vaginalis* can outcompete beneficial microorganisms there and infect the vagina. (*Trichomonas* infections also can occur in the urethra of males, though often without symptoms.) *T. vaginalis* has a gene that allows it to feed on the vaginal lining, promoting infection. Studies suggest that the protist acquired this gene by horizontal gene transfer from bacterial parasites in the vagina.

➤ Figure 28.5 The parabasalid parasite *Trichomonas vaginalis* (colorized SEM).



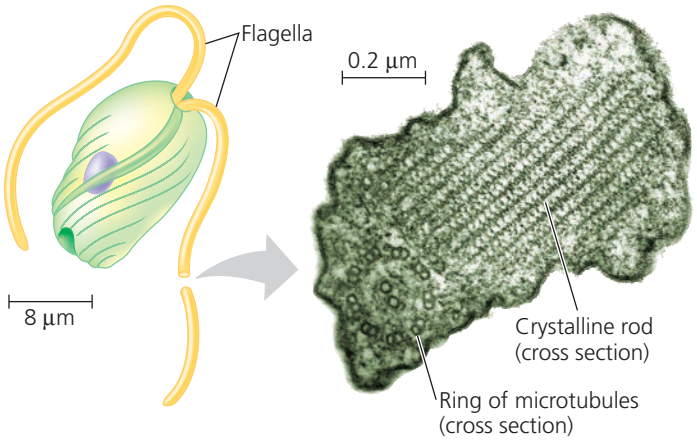


Figure 28.6 Euglenozoan flagellum. Most euglenozoans have a crystalline rod inside one of their flagella (the TEM is a flagellum shown in cross section). The rod lies alongside the 9 + 2 ring of microtubules found in all eukaryotic flagella (compare with Figure 7.24).

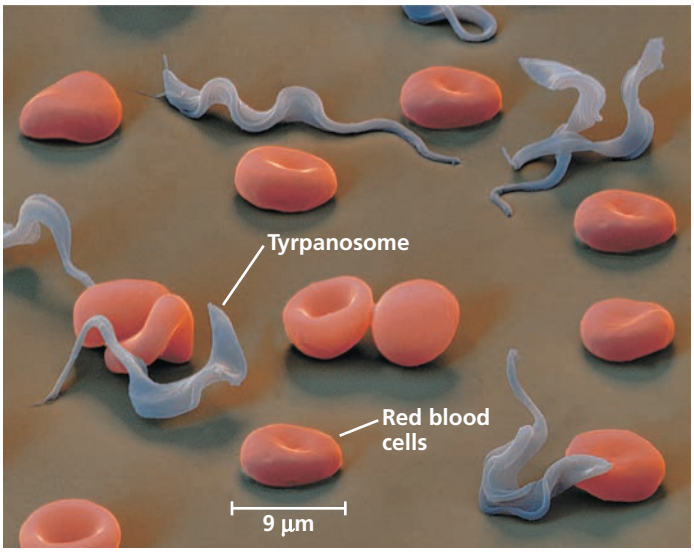


Figure 28.7 *Trypanosoma*, the kinetoplastid that causes sleeping sickness (colorized SEM).

Euglenozoans

Protists called **euglenozoans** belong to a diverse clade that includes predatory heterotrophs, photosynthetic autotrophs, mixotrophs, and parasites. The main morphological feature that distinguishes protists in this clade is the presence of a rod with either a spiral or a crystalline structure inside each of their flagella **(Figure 28.6)**. The two best-studied groups of euglenozoans are the kinetoplastids and the euglenids.

Kinetoplastids

Protists called **kinetoplastids** have a single, large mitochondrion that contains an organized mass of DNA called a *kinetoplast*. These protists include species that feed on prokaryotes in freshwater, marine, and moist terrestrial ecosystems, as well as species that parasitize animals, plants, and other protists. For example, kinetoplastids in the genus *Trypanosoma* infect humans and cause sleeping sickness, a neurological disease that is invariably fatal if not treated. The infection occurs via the bite of a vector (carrier) organism, the African tsetse fly **(Figure 28.7)**. Trypanosomes also cause Chagas' disease, which is transmitted by bloodsucking insects and can lead to congestive heart failure.

Trypanosomes evade immune responses with an effective "bait-and-switch" defense. The surface of a trypanosome is coated with millions of copies of a single protein. However, before the host's immune system can recognize the protein and mount an attack, new generations of the parasite switch to another surface protein with a different molecular structure. Frequent changes in the surface protein prevent the host from developing immunity. (The Scientific Skills Exercise in Chapter 47 explores this topic further.) About a third of *Trypanosoma*'s genome is dedicated to producing these surface proteins.

Euglenids

A **euglenid** has a pocket at one end of the cell from which one or two flagella emerge **(Figure 28.8)**. Some euglenids are

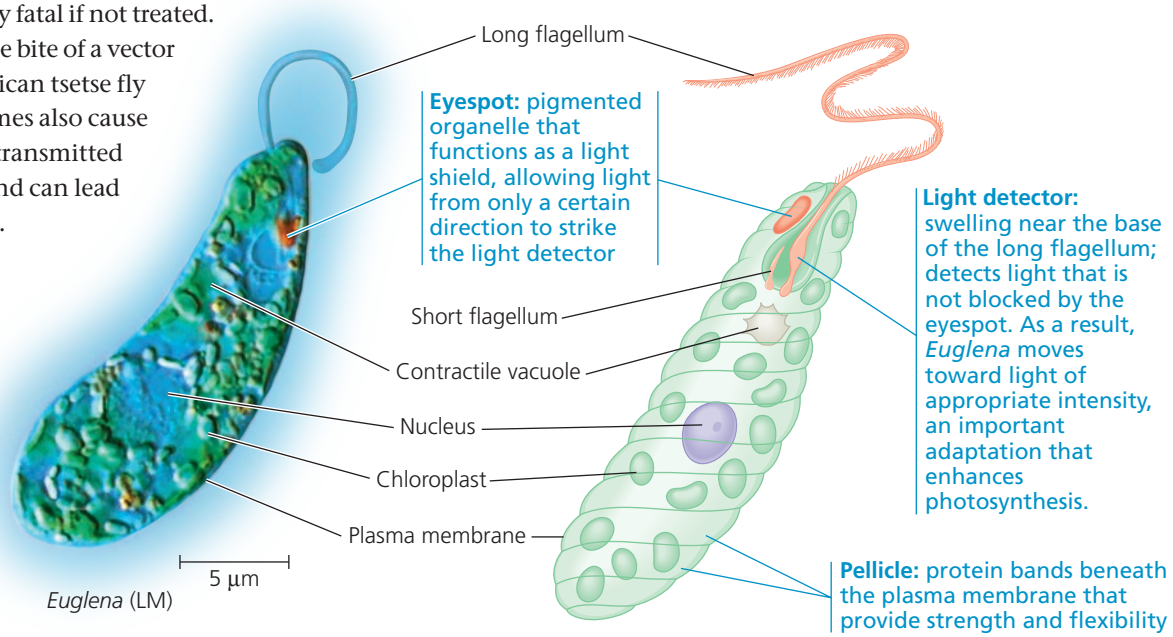


Figure 28.8 *Euglena*, a euglenid commonly found in pond water.

MB **Video: *Euglena***

mixotrophs: They perform photosynthesis when sunlight is available, but when it is not, they can become heterotrophic, absorbing organic nutrients from their environment. Many other euglenids engulf prey by phagocytosis.

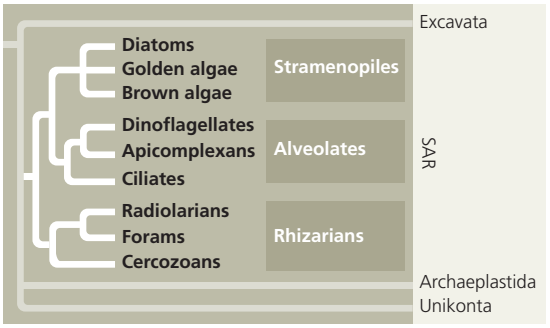
CONCEPT CHECK 28.2

1. Why is it difficult for humans to develop immunity against trypanosomes?
2. **WHAT IF?** ➤ DNA sequence data for a diplomonad, a euglenid, a plant, and an unidentified protist suggest that the unidentified species is most closely related to the diplomonad. Further studies reveal that the unknown species has fully functional mitochondria. Based on these data, at what point on the phylogenetic tree in Figure 28.2 did the mystery protist's lineage probably diverge from other eukaryote lineages? Explain.

For suggested answers, see Appendix A.

CONCEPT 28.3

SAR is a highly diverse group of protists defined by DNA similarities



Our second supergroup, referred to as **SAR**, was proposed recently based on whole-genome DNA sequence analyses. These studies have found that three major clades of protists—the stramenopiles, alveolates, and rhizarians—form a monophyletic supergroup. This supergroup contains a large, extremely diverse collection of protists. To date, this supergroup has not received a formal name but is instead known by the first letters of its major clades: SAR.

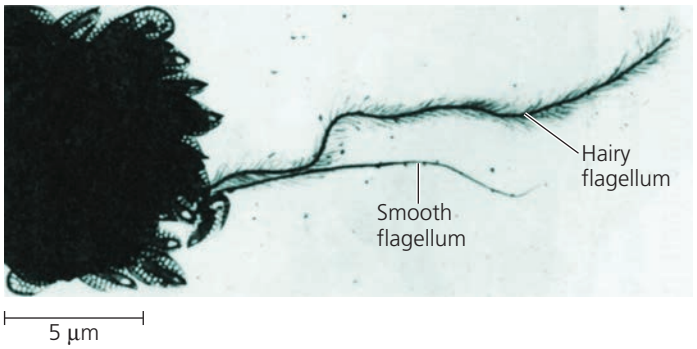
Some morphological and DNA sequence data suggest that two of these groups, the stramenopiles and alveolates, originated more than a billion years ago, when a common ancestor of these two clades engulfed a single-celled, photosynthetic red alga. Because red algae are thought to have originated by primary endosymbiosis (see Figure 28.3), such an origin for the stramenopiles and alveolates is referred to as secondary endosymbiosis. Others question this idea, noting that some species in these groups lack plastids or their remnants (including any trace of plastid genes in their nuclear DNA).

As its lack of a formal name suggests, SAR is one of the most controversial of the four supergroups we describe in this chapter. Even so, for many scientists, this supergroup represents the best current hypothesis for the phylogeny of the three large protist clades to which we now turn.

Stramenopiles

One major subgroup of SAR, the **stramenopiles**, includes some of the most important photosynthetic organisms on the planet. Their name (from the Latin *stramen*, straw, and *pilos*, hair) refers to their characteristic flagellum, which has numerous fine, hairlike projections. In most stramenopiles, this "hairy" flagellum is paired with a shorter "smooth" (nonhairy) flagellum **(Figure 28.9)**. Here we'll focus on three groups of stramenopiles: diatoms, golden algae, and brown algae.

Figure 28.9 Stramenopile flagella. Most stramenopiles, such as *Synura petersenii*, have two flagella: one covered with fine, stiff hairs and a shorter one that is smooth.



Diatoms

A key group of photosynthetic protists, **diatoms** are unicellular algae that have a unique glass-like wall made of silicon dioxide embedded in an organic matrix **(Figure 28.10)**. The wall consists of two parts that overlap like a shoe box and its lid. These walls provide effective protection from the crushing jaws of predators: Live diatoms can withstand pressures as great as 1.4 million kg/m^2, equal to the pressure under each leg of a table supporting an elephant!

With an estimated 100,000 living species, diatoms are a highly diverse group of protists (see Figure 28.2). They are among the most abundant photosynthetic organisms both in the ocean and in lakes: One bucket of water scooped from the surface of the sea may contain millions of these microscopic algae. The abundance of diatoms in the past is also evident in

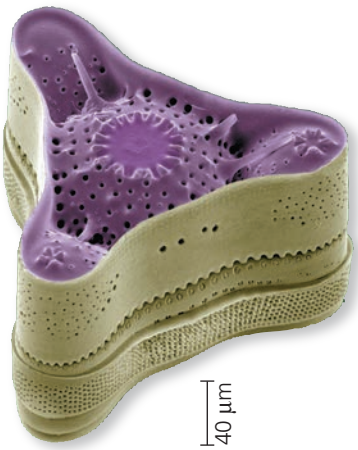


Figure 28.10 The diatom *Triceratium morlandii* (colorized SEM).

the fossil record, where massive accumulations of fossilized diatom walls are major constituents of sediments known as *diatomaceous earth*. These sediments are mined for their quality as a filtering medium and for many other uses.

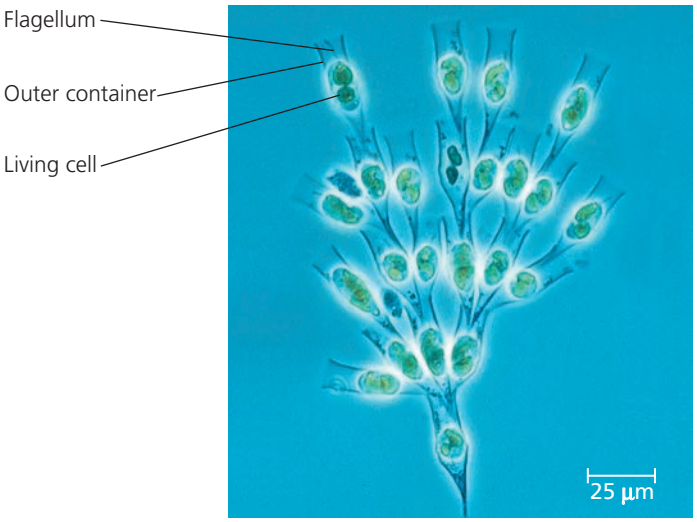
Diatoms are so widespread and abundant that their photosynthetic activity affects global carbon dioxide (CO_2) levels. Diatoms have this effect in part because of events that occur during episodes of rapid population growth, or *blooms*, when ample nutrients are available. Typically, diatoms are eaten by a variety of protists and invertebrates, but during a bloom, many escape this fate. When these uneaten diatoms die, their bodies sink to the ocean floor. It takes decades, or even centuries, for diatoms that sink to the ocean floor to be broken down by bacteria and other decomposers. As a result, the carbon in their bodies remains there for some time, rather than being released immediately as CO_2 as the decomposers respire. The overall effect of these events is that CO_2 absorbed by diatoms during photosynthesis is transported, or "pumped," to the ocean floor.

With an eye toward reducing global warming by lowering atmospheric CO_2 levels, some scientists advocate promoting diatom blooms by fertilizing the ocean with essential nutrients such as iron. In a 2012 study, researchers found that CO_2 was indeed pumped to the ocean floor after iron was added to a small region of the ocean. Further tests are planned to examine whether iron fertilization has undesirable side effects (such as oxygen depletion or the production of nitrous oxide, a more potent greenhouse gas than CO_2).

Golden Algae

The characteristic color of **golden algae** results from their yellow and brown carotenoids. The cells of golden algae are typically biflagellated, with both flagella attached near one end of the cell. Most species are unicellular, but some are colonial **(Figure 28.11)**.

Figure 28.11 *Dinobryon*, a colonial golden alga found in fresh water (LM).



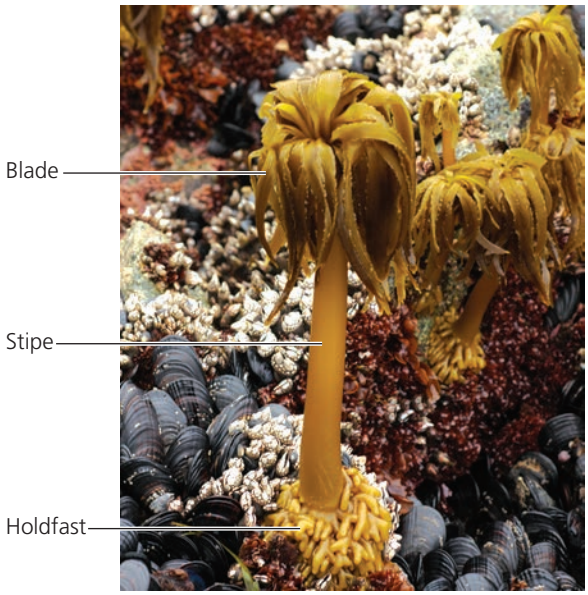
Many golden algae are components of freshwater and marine *plankton*, communities of mostly microscopic organisms that drift in currents near the water's surface. While all golden algae are photosynthetic, some species are mixotrophic. These mixotrophs can absorb dissolved organic compounds or ingest food particles, including living cells, by phagocytosis. If environmental conditions deteriorate, many species form protective cysts that can survive for decades.

Brown Algae

The largest and most complex algae are **brown algae**. All are multicellular, and most are marine. Brown algae are especially common along temperate coasts that have cold-water currents. They owe their characteristic brown or olive color to the carotenoids in their plastids.

Many of the species commonly called "seaweeds" are brown algae. Some brown algal seaweeds have specialized structures that resemble organs in plants, such as a rootlike **holdfast**, which anchors the alga, and a stemlike **stipe**, which supports the leaflike **blades (Figure 28.12)**. Unlike plants, however, brown algae lack true tissues and organs. Moreover, morphological and DNA data show that these similarities evolved independently in the algal and plant lineages and are thus analogous, not homologous. In addition, while plants have adaptations (such as rigid stems) that provide support against gravity, brown algae have adaptations that enable their main photosynthetic surfaces (the leaflike blades) to be near the water surface. Some brown algae accomplish this task with gas-filled, bubble-shaped floats.

Figure 28.12 Seaweeds: adapted to life at the ocean's margins. The sea palm (*Postelsia*) lives on rocks along the coast of the northwestern United States and western Canada. The body of this brown alga is well adapted to maintaining a firm foothold despite the crashing surf.



Giant brown algae known as kelps that live in deep waters have such floats in their blades, which are attached to stipes that can rise as much as 60 m from the seafloor—more than half the length of a football field.

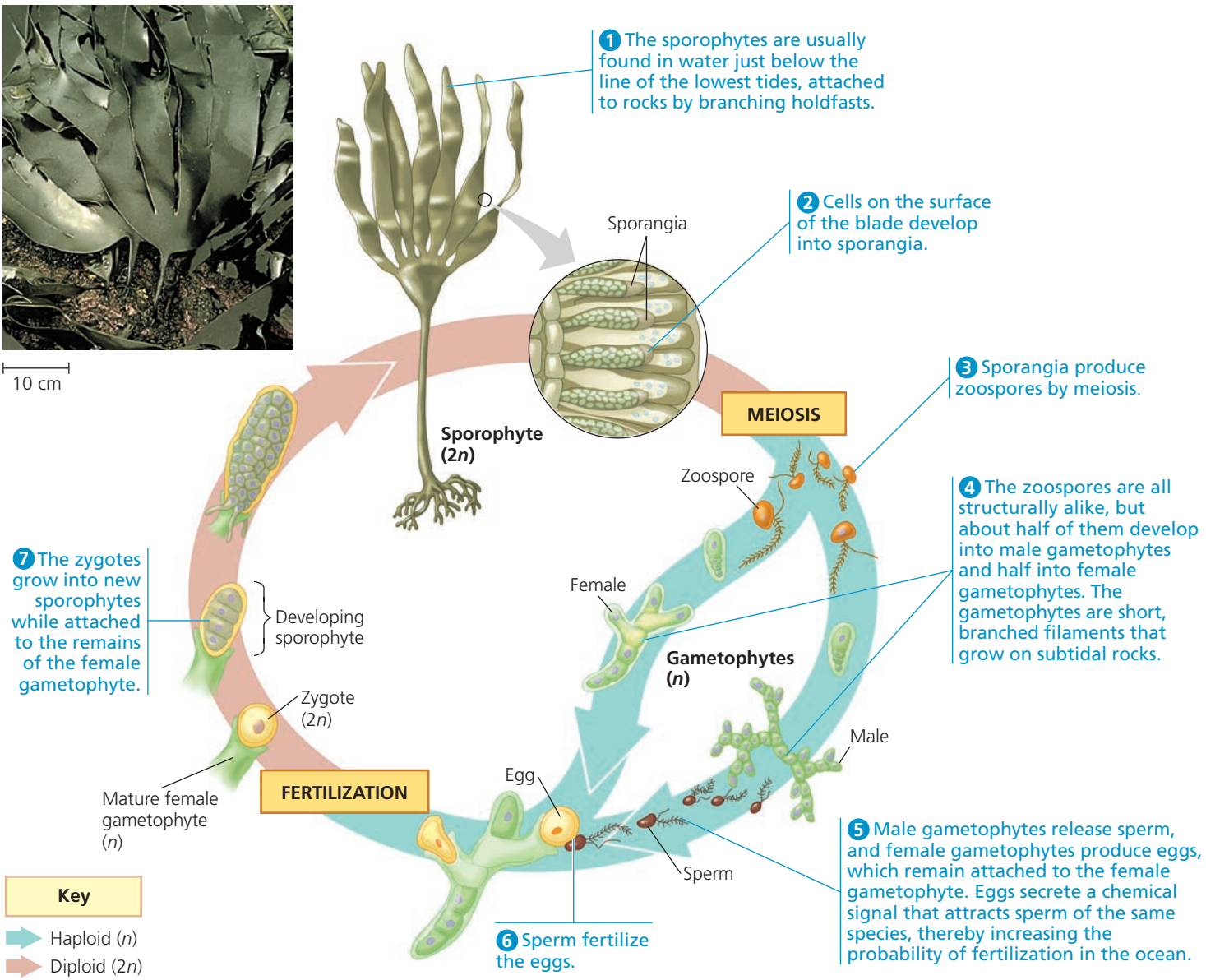
Brown algae are important commodities for humans. Some species are eaten, such as *Laminaria* (Japanese "kombu"), which is used in soups. In addition, the cell walls of brown algae contain a gel-forming substance, called algin, which is used to thicken many processed foods, including pudding and salad dressing.

Alternation of Generations

A variety of life cycles have evolved among the multicellular algae. The most complex life cycles include an **alternation of generations**, the alternation of multicellular haploid and diploid forms. Although haploid and diploid conditions alternate in *all* sexual life cycles—human gametes, for example, are haploid—the term *alternation of generations* applies only to life cycles in which both haploid and diploid stages are multicellular. As you will read in Concept 29.1, alternation of generations also evolved in plants.

The complex life cycle of the brown alga *Laminaria* provides an example of alternation of generations **(Figure 28.13)**. The diploid individual is called the *sporophyte* because it produces spores. The spores are haploid and move by means of flagella; they are called zoospores. The zoospores develop into haploid, multicellular male and female *gametophytes*, which produce gametes. The union of two gametes (fertilization)

Figure 28.13 The life cycle of the brown alga *Laminaria*: an example of alternation of generations.



VISUAL SKILLS ➤ *Based on this diagram, are the sperm shown in* 5 *genetically identical to one another? Explain.*

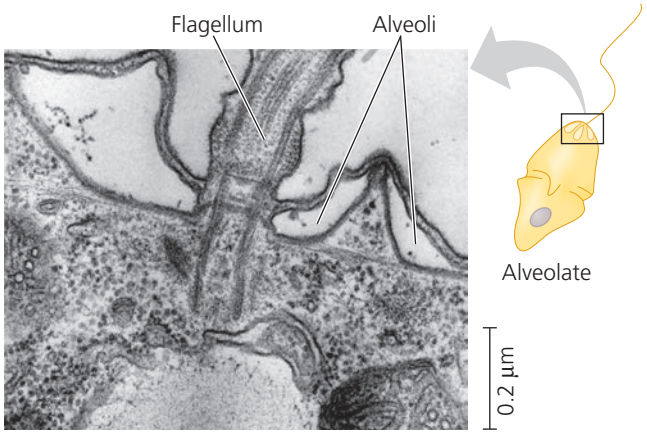
results in a diploid zygote, which matures and gives rise to a new multicellular sporophyte.

In *Laminaria*, the two generations are **heteromorphic**, meaning that the sporophytes and gametophytes are structurally different. Other algal life cycles have an alternation of **isomorphic** generations, in which the sporophytes and gametophytes look similar to each other, although they differ in chromosome number.

Alveolates

Members of the next subgroup of SAR, the **alveolates**, have membrane-enclosed sacs (alveoli) just under the plasma membrane **(Figure 28.14)**. Alveolates are abundant in many habitats and include a wide range of photosynthetic and heterotrophic protists. We'll discuss three alveolate clades here: a group of flagellates (the dinoflagellates), a group of parasites (the apicomplexans), and a group of protists that move using cilia (the ciliates).

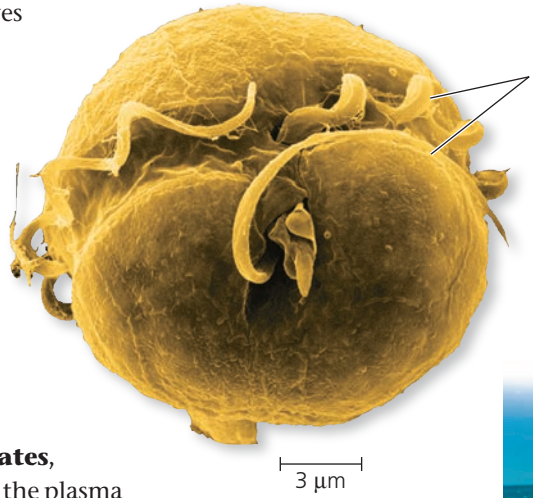
Figure 28.14 Alveoli. These sacs under the plasma membrane are a characteristic that distinguishes alveolates from other eukaryotes (TEM).



Dinoflagellates

The cells of many **dinoflagellates** are reinforced by cellulose plates. Two flagella located in grooves in this "armor" make dinoflagellates (from the Greek *dinos*, whirling) spin as they move through the waters of their marine and freshwater communities **(Figure 28.15a)**. Although their ancestors may have originated by secondary endosymbiosis (see Figure 28.3), roughly half of all dinoflagellates are now purely heterotrophic. Others are important species of *phytoplankton* (photosynthetic plankton, which include photosynthetic bacteria as well as algae); many photosynthetic dinoflagellates are mixotrophic.

Periods of explosive population growth (blooms) in dinoflagellates sometimes cause a phenomenon called "red tide"



(a) Dinoflagellate flagella. Beating of the spiral flagellum, which lies in a groove that encircles the cell, makes this specimen of *Pfiesteria shumwayae* spin (colorized SEM).

(b) Red tide in the Gulf of Carpentaria in northern Australia. The red color is due to high concentrations of a carotenoid-containing dinoflagellate.

Figure 28.15 Dinoflagellates.

MB Video: Dinoflagellate

(Figure 28.15b). The blooms make coastal waters appear brownish red or pink because of the presence of carotenoids, the most common pigments in dinoflagellate plastids. Toxins produced by certain dinoflagellates have caused massive kills of invertebrates and fishes. Humans who eat molluscs that have accumulated the toxins are affected as well, sometimes fatally.

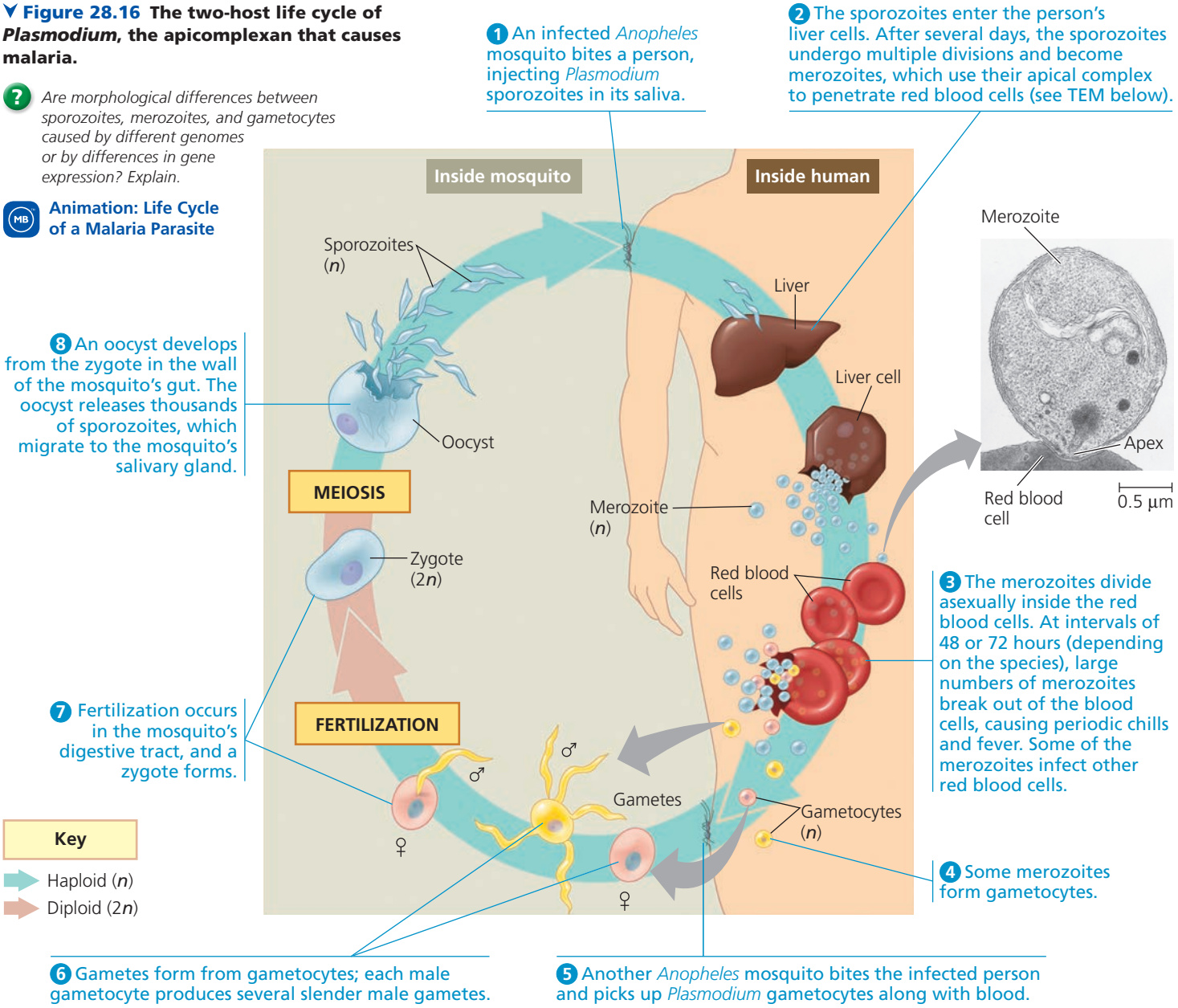
Apicomplexans

Nearly all **apicomplexans** are parasites of animals—and virtually all animal species examined so far are attacked by these parasites. The parasites spread through their host as tiny infectious cells called *sporozoites*. Apicomplexans are so named because one end (the *apex*) of the sporozoite cell contains a *complex* of organelles specialized for penetrating host cells and tissues. Although apicomplexans are not photosynthetic, recent data show that they retain a modified plastid (apicoplast), most likely of red algal origin.

Figure 28.16 The two-host life cycle of *Plasmodium*, the apicomplexan that causes malaria.

? *Are morphological differences between sporozoites, merozoites, and gametocytes caused by different genomes or by differences in gene expression? Explain.*

MB **Animation: Life Cycle of a Malaria Parasite**



Most apicomplexans have intricate life cycles with both sexual and asexual stages. Those life cycles often require two or more host species for completion. For example, *Plasmodium*, the parasite that causes malaria, lives in both mosquitoes and humans **(Figure 28.16)**.

Historically, malaria has rivaled tuberculosis as the leading cause of human death by infectious disease. The incidence of malaria was diminished in the 1960s by insecticides that reduced carrier populations of *Anopheles* mosquitoes and by drugs that killed *Plasmodium* in humans. But the emergence of resistant varieties of both *Anopheles* and *Plasmodium* has led to a resurgence of malaria. About 200 million people in the tropics are currently infected, and 600,000 die each year. In regions where malaria is common, the lethal effects of this disease have resulted in the evolution of high frequencies of the sickle-cell allele; for an explanation of this connection, see Figure 23.18.

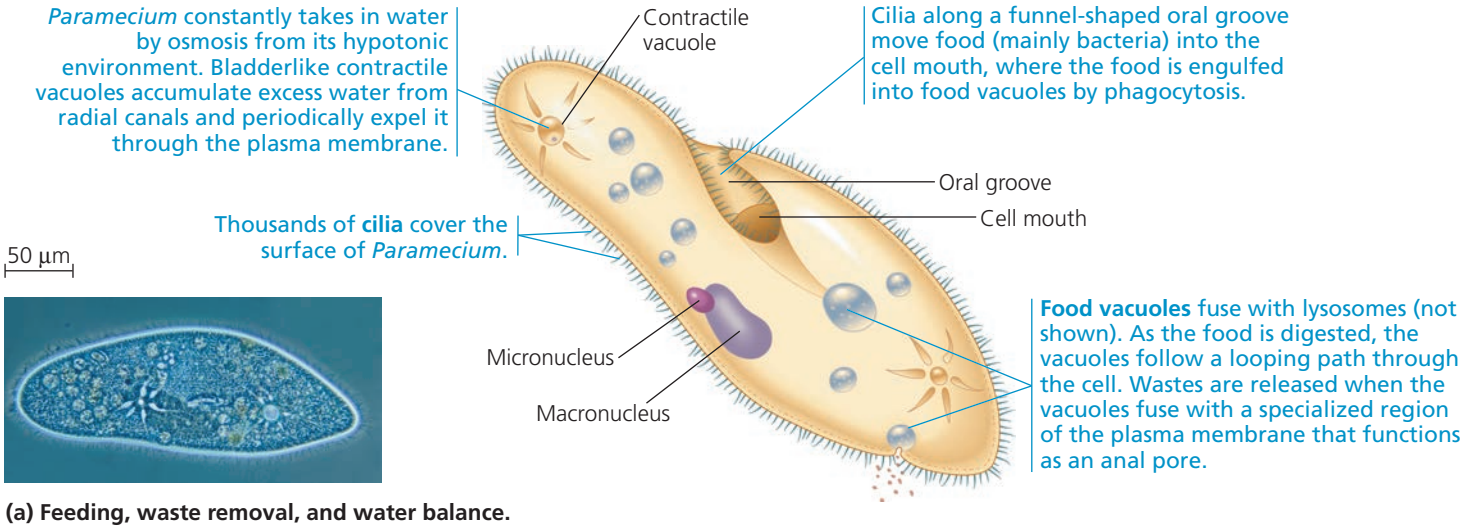
The search for malarial vaccines has been hampered by the fact that *Plasmodium* lives mainly inside cells, hidden from the host's immune system. And, like trypanosomes, *Plasmodium* continually changes its surface proteins. Even so, significant progress was made in 2015, when European regulators approved the world's first licensed malarial vaccine. However, this vaccine, which targets a protein on the surface of sporozoites, provides only partial protection against malaria. As a result, researchers continue to study other potential vaccine targets, including the apicoplast. This approach may be effective because the apicoplast is a modified plastid; as such, it descended from a cyanobacterium and hence has different metabolic pathways from those in the human patients.

Ciliates

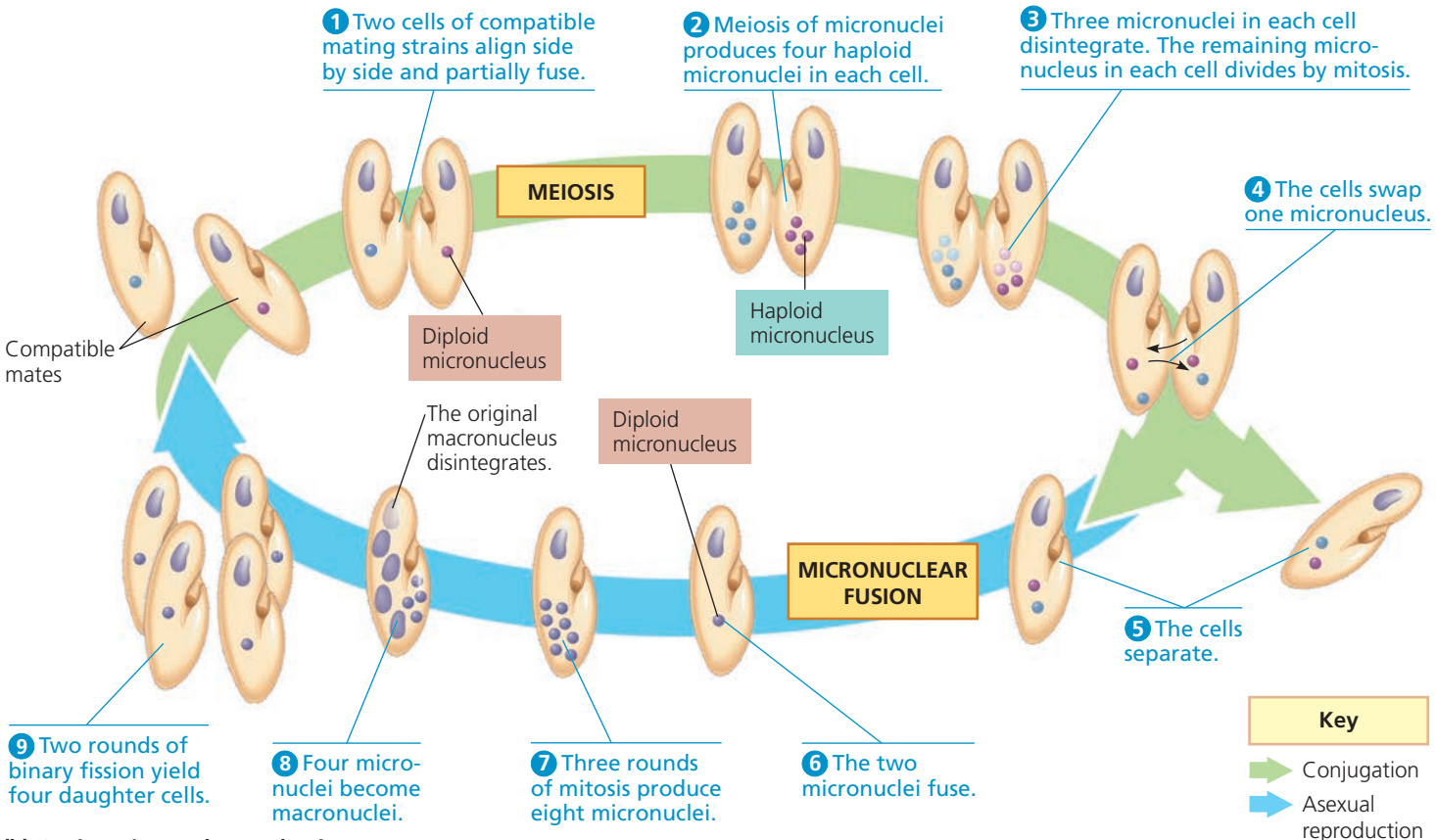
The **ciliates** are a large and varied group of protists named for their use of cilia to move and feed **(Figure 28.17a)**. Most ciliates are predators, typically of bacteria or of other protists. Their cilia may completely cover the cell surface or may be clustered in a few rows or tufts. In certain species, rows of tightly packed cilia function collectively in locomotion. Other ciliates scurry about on leg-like structures constructed from many cilia bonded together.

A distinctive feature of ciliates is the presence of two types of nuclei: tiny micronuclei and large macronuclei. A cell has one or more nuclei of each type. Genetic variation

Figure 28.17 Structure and function in the ciliate *Paramecium caudatum*.



(a) Feeding, waste removal, and water balance.



(b) Conjugation and reproduction.

MAKE CONNECTIONS ➤ *The events shown in steps 5 and 6 of this diagram have a similar overall effect to what event in the human life cycle (see Figure 13.5)? Explain.*

Video: *Paramecium*

results from **conjugation**, a sexual process in which two individuals exchange haploid micronuclei but do not reproduce **(Figure 28.17b)**. Ciliates generally reproduce asexually by binary fission, during which the existing macronucleus disintegrates and a new one is formed from the cell's micronuclei. Each macronucleus typically contains multiple copies of the ciliate's genome. Genes in the macronucleus control the everyday functions of the cell, such as feeding, waste removal, and maintaining water balance.

MB Video: Ciliate Movement in *Stentor*

Rhizarians

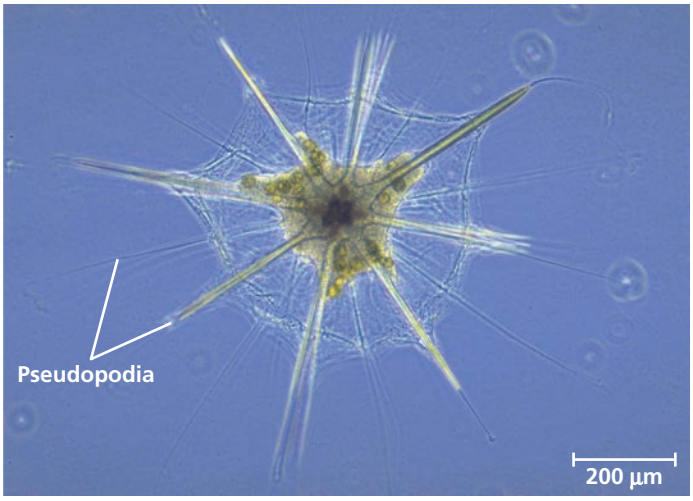
Our next subgroup of SAR is the **rhizarians**. Many species in this group are **amoebas**, protists that move and feed by means of **pseudopodia**, extensions that may bulge from almost anywhere on the cell surface. As it moves, an amoeba extends a pseudopodium and anchors the tip; more cytoplasm then streams into the pseudopodium. Amoebas do not constitute a monophyletic group; instead, they are dispersed across many distantly related eukaryotic taxa. Most amoebas that are rhizarians differ morphologically from other amoebas by having threadlike pseudopodia. Rhizarians also include flagellated (non-amoeboid) protists that feed using threadlike pseudopodia.

We'll examine three groups of rhizarians here: radiolarians, forams, and cercozoans.

Radiolarians

The protists called **radiolarians** have delicate, intricately symmetrical internal skeletons that are generally made of silica. The pseudopodia of these mostly marine protists radiate from the central body **(Figure 28.18)** and are reinforced by bundles of microtubules. The microtubules are covered by a thin layer of cytoplasm, which engulfs smaller microorganisms that become attached to the pseudopodia. Cytoplasmic streaming then carries the captured prey into the main part of the cell. After radiolarians die, their skeletons settle to the seafloor, where they have accumulated as an ooze that is hundreds of meters thick in some locations.

Figure 28.18 A radiolarian. Numerous threadlike pseudopodia radiate from the central body of this radiolarian (LM).



Forams

The protists called **foraminiferans** (from the Latin *foramen*, little hole, and *ferre*, to bear), or **forams**, are named for their porous shells, called **tests** (see Figure 28.2). Foram tests consist of a single piece of organic material that typically is hardened with calcium carbonate. The pseudopodia that extend through the pores function in swimming, test formation, and feeding. Many forams also derive nourishment from the photosynthesis of symbiotic algae that live within the tests.

Forams are found in both the ocean and fresh water. Most species live in sand or attach themselves to rocks or algae, but some live as plankton. The largest forams, though single-celled, have tests several centimeters in diameter.

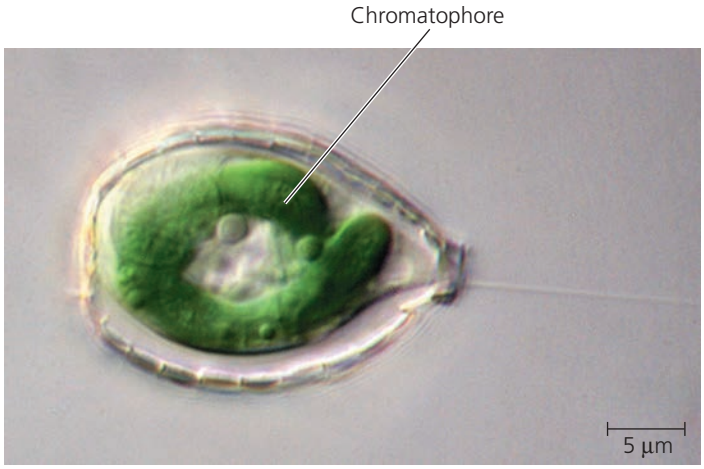
Ninety percent of all identified species of forams are known from fossils. Along with the calcium-containing remains of other protists, the fossilized tests of forams are part of marine sediments, including sedimentary rocks that are now land formations. Foram fossils are excellent markers for correlating the ages of sedimentary rocks in different parts of the world. Researchers are also studying these fossils to obtain information about climate change and its effects on the oceans and their life **(Figure 28.19)**.

Cercozoans

First identified in molecular phylogenies, the **cercozoans** are a large group of amoeboid and flagellated protists that feed using threadlike pseudopodia. Cercozoan protists are common inhabitants of marine, freshwater, and soil ecosystems.

Figure 28.19 Fossil forams. By measuring the magnesium content in fossilized forams like these, researchers seek to learn how ocean temperatures have changed over time. Forams take up more magnesium in warmer water than in colder water.

Figure 28.20 A second case of primary endosymbiosis? The cercozoan *Paulinella* conducts photosynthesis in a unique sausage-shaped structure called a chromatophore (LM). Chromatophores are surrounded by a membrane with a peptidoglycan layer, suggesting that they are derived from a bacterium. DNA evidence indicates that chromatophores are derived from a different cyanobacterium than that from which plastids are derived.



Most cercozoans are heterotrophs. Many are parasites of plants, animals, or other protists; many others are predators. The predators include the most important consumers of bacteria in aquatic and soil ecosystems, along with species that eat other protists, fungi, and even small animals. One small group of cercozoans, the chlorarachniophytes (mentioned earlier in the discussion of secondary endosymbiosis), are mixotrophic: These organisms ingest smaller protists and bacteria as well as perform photosynthesis. At least one other cercozoan, *Paulinella chromatophora*, is an autotroph, deriving its energy from light and its carbon from CO_2. As described in **Figure 28.20**, *Paulinella* appears to represent an intriguing additional evolutionary example of a eukaryotic lineage that obtained its photosynthetic apparatus directly from a cyanobacterium.

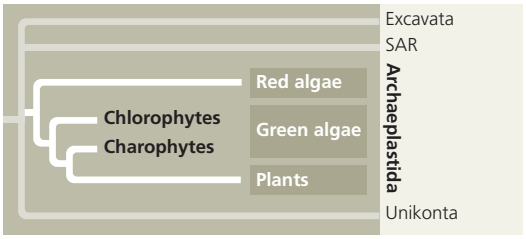
CONCEPT CHECK 28.3

1. Why is it difficult to develop a vaccine against malaria?
2. **WHAT IF? ➤** Would you expect the plastid DNA of photosynthetic dinoflagellates, diatoms, and golden algae to be more similar to the nuclear DNA of plants (domain Eukarya) or to the chromosomal DNA of cyanobacteria (domain Bacteria)? Explain.
3. **MAKE CONNECTIONS ➤** Which of the three life cycles in Figure 13.6 exhibits alternation of generations? How does it differ from the other two?
4. **MAKE CONNECTIONS ➤** Review Figures 10.2 and 11.6, and then summarize how CO_2 and O_2 are both used and produced by chlorarachniophytes and other aerobic algae.

For suggested answers, see Appendix A.

CONCEPT 28.4

Red algae and green algae are the closest relatives of plants



As described earlier, morphological and molecular evidence indicates that plastids arose when a heterotrophic protist acquired a cyanobacterial endosymbiont. Later, photosynthetic descendants of this ancient protist evolved into red algae and green algae (see Figure 28.3), and the lineage that produced green algae then gave rise to plants. Together, red algae, green algae, and plants make up our third eukaryotic supergroup, which is called **Archaeplastida**. Archaeplastida is a monophyletic group that descended from the ancient protist that engulfed a cyanobacterium. We will examine plants in Chapters 29 and 30; here we will look at the diversity of their closest algal relatives, red algae and green algae.

Red Algae

Many of the 6,000 known species of **red algae** (rhodophytes, from the Greek *rhodos*, red) are reddish, owing to the photosynthetic pigment phycoerythrin, which masks the green of chlorophyll **(Figure 28.21)**. However, other species (those adapted to shallow water) have less phycoerythrin. As a result, red algal species may be greenish red in very shallow water, bright red at moderate depths, and almost black in deep water. Some species lack pigmentation altogether and live as heteroptrophic parasites on other red algae.

Red algae are abundant in the warm coastal waters of tropical oceans. Some of their photosynthetic pigments, including phycoerythrin, allow them to absorb blue and green light, which penetrate relatively far into the water. A species of red alga has been discovered near the Bahamas at a depth of more than 260 m. There are also a small number of freshwater and terrestrial species.

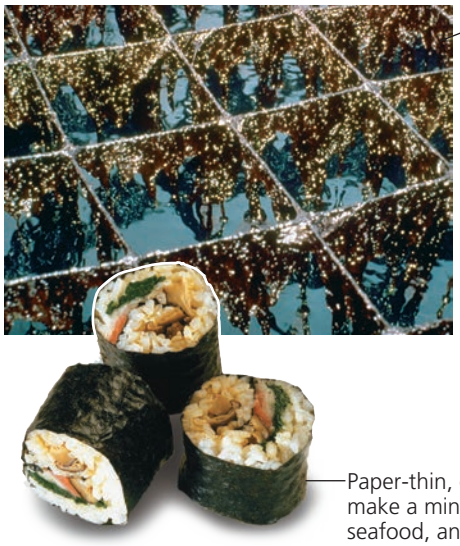
Most red algae are multicellular. Although none are as big as the giant brown kelps, the largest multicellular red algae are included in the informal designation "seaweeds." You may have eaten one of these multicellular red algae, *Porphyra* (Japanese "nori"), as crispy sheets or as a wrap for sushi (see Figure 28.21). Red algae reproduce sexually and have diverse life cycles in which alternation of generations is common. However, unlike other algae, red algae do not have flagellated gametes, so they depend on water currents to bring gametes together for fertilization.

Figure 28.21 Red algae.

▶ ***Bonnemaisonia hamifera*.** This red alga has a filamentous form.

◀ **Dulse (*Palmaria palmata*).** This edible species has a "leafy" form.

▼ **Nori.** The red alga *Porphyra* is the source of a traditional Japanese food.



The seaweed is grown on nets in shallow coastal waters.

Paper-thin, glossy sheets of dried nori make a mineral-rich wrap for rice, seafood, and vegetables in sushi.

Green Algae

The grass-green chloroplasts of **green algae** have a structure and pigment composition much like the chloroplasts of plants. Molecular systematics and cellular morphology leave little doubt that green algae and plants are closely related. In fact, some systematists now advocate including green algae in an expanded "plant" kingdom, Viridiplantae (from the Latin *viridis*, green). Phylogenetically, this change makes sense, since otherwise the green algae are a paraphyletic group.

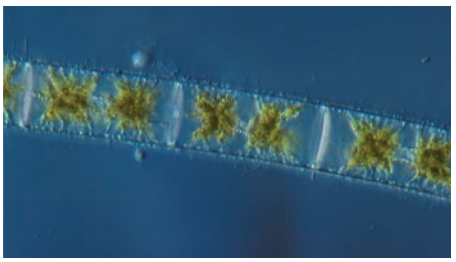
Green algae can be divided into two main groups, the charophytes and the chlorophytes. The charophytes include the algae most closely related to plants, and we will discuss them along with plants in Chapter 29.

The second group, the chlorophytes (from the Greek *chloros*, green), includes more than 7,000 species. Most live in fresh water, but there are also many marine and some terrestrial species. The simplest chlorophytes are unicellular organisms such as *Chlamydomonas*, which resemble gametes of more complex chlorophytes. Various species of unicellular chlorophytes live independently in aquatic habitats as phytoplankton or inhabit damp soil. Some live symbiotically within other eukaryotes, contributing part of their photosynthetic output to the food supply of their hosts. Still other chlorophytes live in environments exposed to intense visible and ultraviolet radiation; these species are protected by radiation-blocking compounds in their cytoplasm, cell wall, or zygote coat.

Larger size and greater complexity evolved in green algae by three different mechanisms:

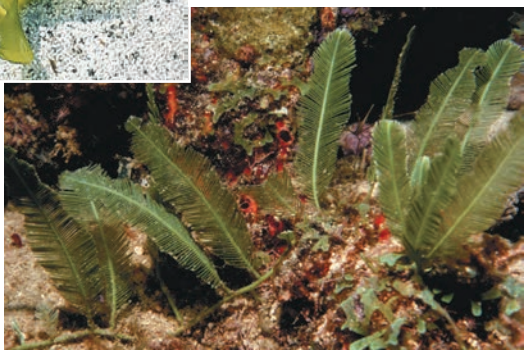
1. The formation of colonies of individual cells, as seen in *Zygnema* **(Figure 28.22a)** and other species whose filamentous forms contribute to the stringy masses known as pond scum.
2. The formation of true multicellular bodies by cell division and differentiation, as in *Volvox* (see Figure 28.2) and *Ulva* **(Figure 28.22b)**.
3. The repeated division of nuclei with no cytoplasmic division, as in *Caulerpa* **(Figure 28.22c)**.

Figure 28.22 Examples of large chlorophytes.



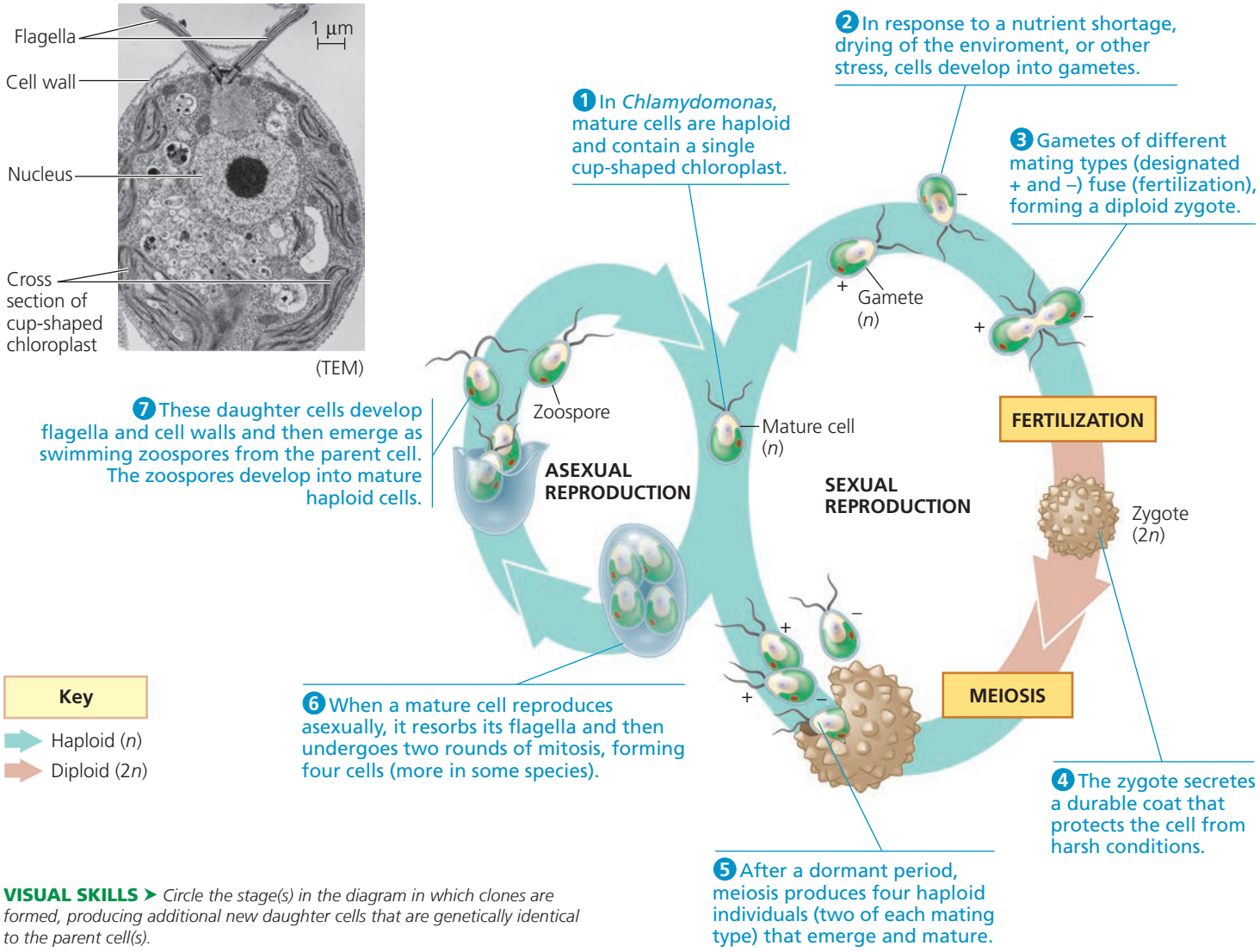
(a) *Zygnema*, a common pond alga. This filamentous charophyte features two star-shaped chloroplasts in each cell.

(b) *Ulva*, or sea lettuce. This multicellular, edible chlorophyte has differentiated structures, such as its leaflike blades and a rootlike holdfast that anchors the alga.



(c) *Caulerpa*, an intertidal chlorophyte. The branched filaments lack crosswalls and thus are multinucleate. In effect, the body of this alga is one huge "supercell."

Figure 28.23 The life cycle of *Chlamydomonas*, a unicellular chlorophyte.



VISUAL SKILLS ➤ *Circle the stage(s) in the diagram in which clones are formed, producing additional new daughter cells that are genetically identical to the parent cell(s).*

Most chlorophytes have complex life cycles, with both sexual and asexual reproductive stages. Nearly all species of chlorophytes reproduce sexually by means of biflagellated gametes that have cup-shaped chloroplasts **(Figure 28.23)**. Alternation of generations has evolved in some chlorophytes, including *Ulva*.

Animation: Alternation of Generations in a Protist

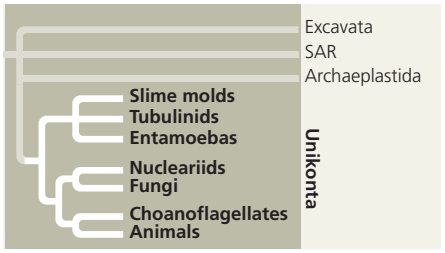
CONCEPT CHECK 28.4

1. Contrast red algae and brown algae.
2. Why is it accurate to say that *Ulva* is truly multicellular but *Caulerpa* is not?
3. **WHAT IF? ➤** Suggest a possible reason why species in the green algal lineage may have been more likely to colonize land than species in the red algal lineage.

For suggested answers, see Appendix A.

CONCEPT 28.5

Unikonts include protists that are closely related to fungi and animals



Unikonta is an extremely diverse supergroup of eukaryotes that includes animals, fungi, and some protists. There are two major clades of unikonts, the amoebozoans and the

opisthokonts (animals, fungi, and closely related protist groups). Each of these two major clades is strongly supported by molecular systematics. The close relationship between amoebozoans and opisthokonts is more controversial. Support for this close relationship is provided by comparisons of myosin proteins and by some (but not all) studies based on multiple genes or whole genomes.

Another controversy involving the unikonts concerns the root of the eukaryotic tree. Recall that the root of a phylogenetic tree anchors the tree in time: Branch points close to the root are the oldest. At present, the root of the eukaryotic tree is uncertain; hence, we do not know which supergroup of eukaryotes was the first to diverge from all other eukaryotes. Some hypotheses, such as the amitochondriate hypothesis described earlier, have been abandoned, but researchers have yet to agree on an alternative. If the root of the eukaryotic tree were known, it would help scientists infer characteristics of the common ancestor of all eukaryotes.

In trying to determine the root of the eukaryotic tree, researchers have based their phylogenies on different sets of genes, some of which have produced conflicting results. Researchers have also tried a different approach, based on tracing the occurrence of a rare evolutionary event **(Figure 28.24)**. Results from this "rare event" approach indicate that Excavata, SAR, and Archaeplastida share a more recent common ancestor than any of them does with Unikonta. This suggests that the root of the tree is located between the unikonts and all other eukaryotes, which implies that the unikonts were the first eukaryotic supergroup to diverge from all other eukaryotes. This idea remains controversial and will require more supporting evidence to be widely accepted.

Amoebozoans

The **amoebozoan** clade includes many species of amoebas that have lobe- or tube-shaped pseudopodia rather than the threadlike pseudopodia found in rhizarians. Amoebozoans include slime molds, tubulinids, and entamoebas.

Slime Molds

Slime molds, or mycetozoans (from the Latin, meaning "fungus animals"), once were thought to be fungi because, like fungi, they produce fruiting bodies that aid in spore dispersal. However, DNA sequence analyses indicate that the resemblance between slime molds and fungi is a case of evolutionary convergence. DNA sequence analyses also show that slime molds descended from unicellular ancestors—an example of the independent origin of multicellularity in eukaryotes.

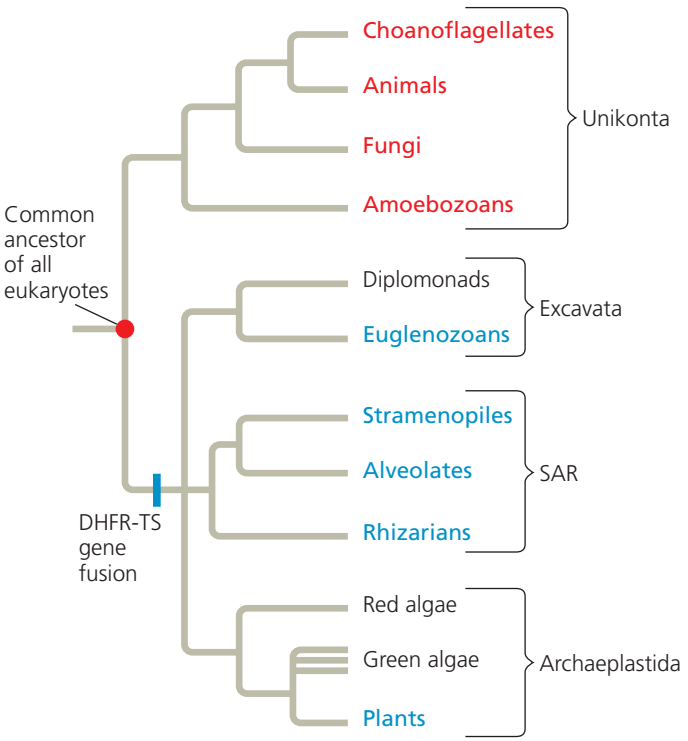
Slime molds have diverged into two main branches, plasmodial slime molds and cellular slime molds. We'll compare their characteristics and life cycles.

Figure 28.24

Inquiry What is the root of the eukaryotic tree?

Experiment Responding to the difficulty in determining the root of the eukaryotic phylogenetic tree, Alexandra Stechmann and Thomas Cavalier-Smith proposed a new approach. They studied two genes, one coding for the enzyme dihydrofolate reductase (DHFR) and the other coding for the enzyme thymidylate synthase (TS). The scientists' approach took advantage of a rare evolutionary event: In some organisms, the genes for DHFR and TS have fused, leading to the production of a single protein with both enzyme activities. Stechmann and Cavalier-Smith amplified (using PCR; see Figure 19.8) and sequenced the genes for DHFR and TS in nine species (one choanoflagellate, two amoebozoans, one euglenozoan, one stramenopile, one alveolate, and three rhizarians). They combined their data with previously published data for species of bacteria, animals, plants, and fungi.

Results The bacteria studied all have separate genes coding for DHFR and TS, suggesting that this is the ancestral condition (red dot on the tree below). Other taxa with separate genes are denoted by red type. Fused genes are a derived character, found in certain members (blue type) of the supergroups Excavata, SAR, and Archaeplastida:

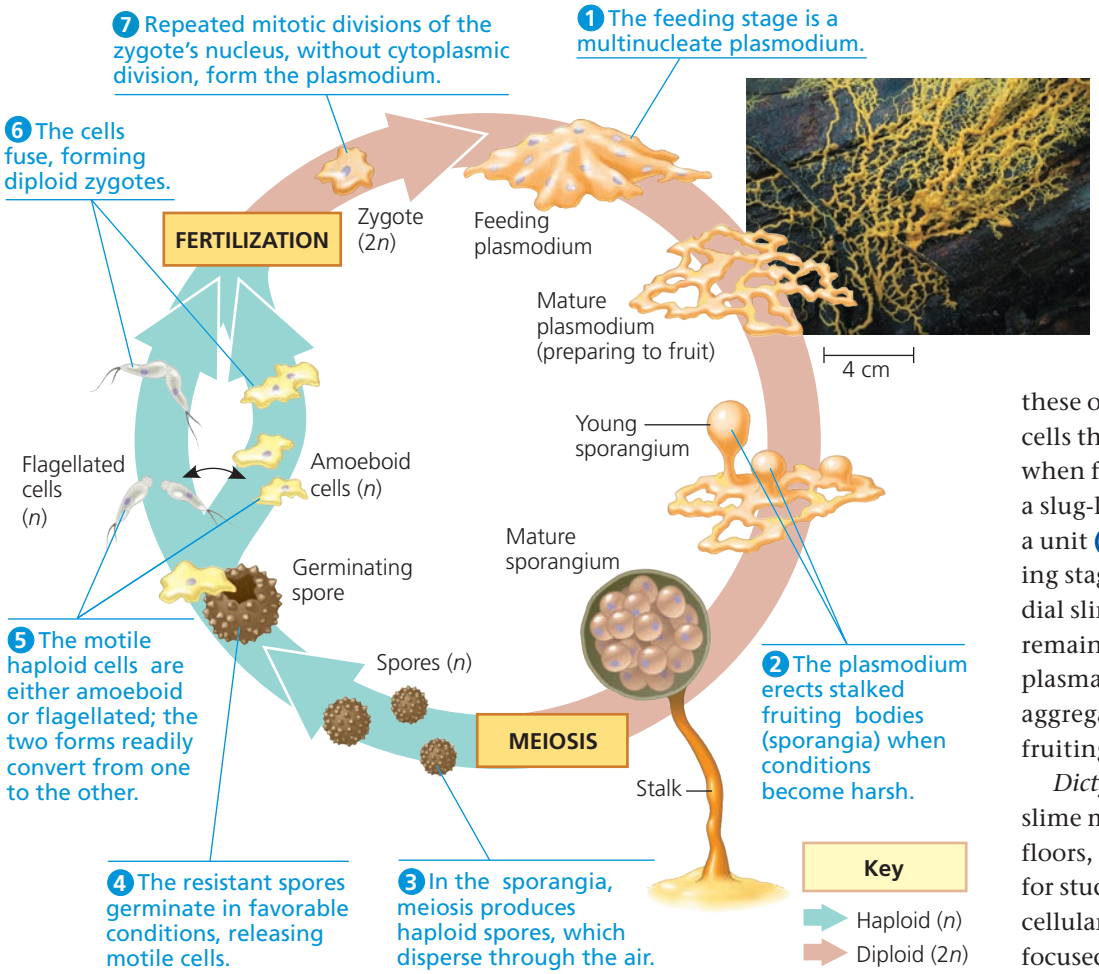


Conclusion The results show that Excavata, SAR, and Archaeplastida form a clade, which supports the hypothesis that the root of the tree is located between the unikonts and all other eukaryotes. Because support for this hypothesis is based on only one trait—the fusion of the genes for DHFR and TS—more data are needed to evaluate its validity.

Data from A. Stechmann and T. Cavalier-Smith, Rooting the eukaryote tree by using a derived gene fusion, *Science* 297:89–91 (2002).

WHAT IF? ➤ *Stechmann and Cavalier-Smith wrote that their conclusions are "valid only if the genes fused just once and were never secondarily split." Why is this assumption critical to their approach?*

Figure 28.25 A plasmodial slime mold. The photograph shows a mature plasmodium, the feeding stage in the life cycle of a plasmodial slime mold. When food becomes scarce, the plasmodium forms stalked fruiting bodies that produce haploid spores that function in sexual reproduction.



Plasmodial Slime Molds Many plasmodial slime molds are brightly colored, often yellow or orange **(Figure 28.25)**. As they grow, they form a mass called a plasmodium, which can be many centimeters in diameter. (Don't confuse a slime mold's plasmodium with the genus *Plasmodium*, which includes the parasitic apicomplexan that causes malaria.) Despite its size, the plasmodium is not multicellular; it is a single mass of cytoplasm that is undivided by plasma membranes and that contains many nuclei. This "supercell" is the product of mitotic nuclear divisions that are not followed by cytokinesis. The plasmodium extends pseudopodia through moist soil, leaf mulch, or rotting logs, engulfing food particles by phagocytosis as it grows. If the habitat begins to dry up or there is no food left, the plasmodium stops growing and differentiates into fruiting bodies that function in sexual reproduction.

Cellular Slime Molds The life cycle of the protists called cellular slime molds can prompt us to question what it means to be an individual organism. The feeding stage of these organisms consists of solitary cells that function individually, but when food is depleted, the cells form a slug-like aggregate that functions as a unit **(Figure 28.26)**. Unlike the feeding stage (plasmodium) of a plasmodial slime mold, these aggregated cells remain separated by their individual plasma membranes. Ultimately, the aggregated cells form an asexual fruiting body.

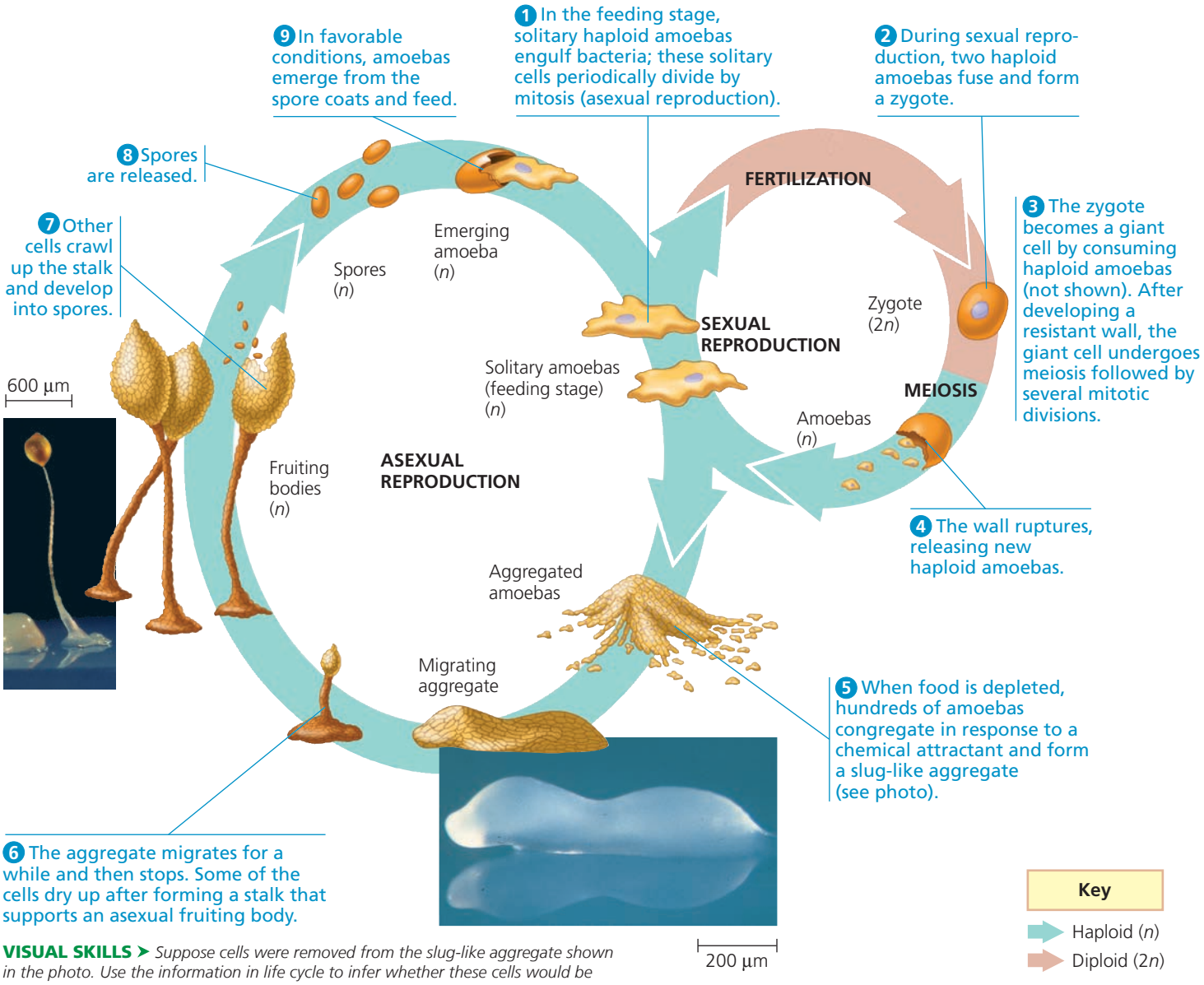
Dictyostelium discoideum, a cellular slime mold commonly found on forest floors, has become a model organism for studying the evolution of multicellularity. One line of research has focused on the slime mold's fruiting body stage. During this stage, the cells that form the stalk die as they dry out, while the spore cells at the top survive and have the potential to reproduce (see Figure 28.26). Scientists have found that mutations in a single gene can turn individual *Dictyostelium* cells into "cheaters" that never become part of the stalk. Because these mutants gain a strong reproductive advantage over noncheaters, why don't all *Dictyostelium* cells cheat?

Recent discoveries suggest an answer to this question. Cheating cells lack a specific surface protein, and noncheating cells can recognize this difference. Noncheaters preferentially aggregate with other noncheaters, thus depriving cheaters of the chance to exploit them. Such a recognition system may have been important in the evolution of other multicellular eukaryotes, such as animals and plants.

Tubulinids

Tubulinids constitute a large and varied group of amoebozoans that have lobe- or tube-shaped pseudopodia. These unicellular protists are ubiquitous in soil as well as freshwater and marine environments. Most are heterotrophs that actively seek and consume bacteria and other protists; one such tubulinid

Figure 28.26 The life cycle of *Dictyostelium*, a cellular slime mold.



VISUAL SKILLS ➤ *Suppose cells were removed from the slug-like aggregate shown in the photo. Use the information in life cycle to infer whether these cells would be haploid or diploid. Explain.*

species, *Amoeba proteus*, is shown in Figure 28.2. Some tubulinids also feed on detritus (nonliving organic matter).

Entamoebas

Whereas most amoebozoans are free-living, those that belong to the genus *Entamoeba* are parasites. They infect all classes of vertebrate animals as well as some invertebrates. Humans are host to at least six species of *Entamoeba*, but only one, *E. histolytica*, is known to be pathogenic. *E. histolytica* causes amebic dysentery and is spread via contaminated drinking water, food, or eating utensils. Responsible for up to 100,000 deaths worldwide every year, the disease is the third-leading cause of death due to eukaryotic parasites, after malaria (see Figure 28.16) and schistosomiasis (see Figure 33.11).

Opisthokonts

Opisthokonts are an extremely diverse group of eukaryotes that includes animals, fungi, and several groups of protists. We will discuss the evolutionary history of fungi and animals in Chapters 31–34. Of the opisthokont protists, we will discuss the nucleariids in Chapter 31 because they are more closely related to fungi than they are to other protists. Similarly, we will discuss choanoflagellates in Chapter 32, since they are more closely related to animals than they are to other protists. The nucleariids and choanoflagellates illustrate why scientists have abandoned the former kingdom Protista: A monophyletic group that includes these single-celled eukaryotes would also have to include the multicellular animals and fungi that are closely related to them.

CONCEPT CHECK 28.5

1. Contrast the pseudopodia of amoebozoans and forams.
2. How does a plasmodial slime mold differ in structure from a cellular slime mold during the feeding stage?
3. **DRAW IT** ➤ Recent evidence indicates that the root of the eukaryotic tree may lie between a clade that includes unikonts and excavates, and all other eukaryotes. Draw the tree suggested by this result.

For suggested answers, see Appendix A.

CONCEPT 28.6

Protists play key roles in ecological communities

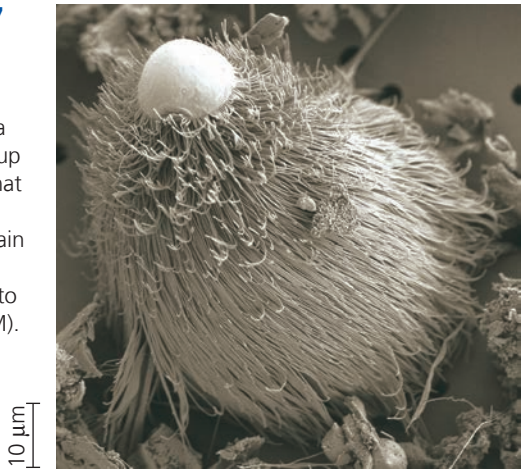
Most protists are aquatic, and they are found almost anywhere there is water, including moist terrestrial habitats such as damp soil and leaf litter. In oceans, ponds, and lakes, many protists are bottom-dwellers that attach to rocks and other substrates or creep through the sand and silt. As we've seen, other protists are important constituents of plankton. We'll focus here on two key roles that protists play in the varied habitats in which they live: that of symbiont and that of producer.

Symbiotic Protists

Many protists form symbiotic associations with other species. For example, photosynthetic dinoflagellates are food-providing symbiotic partners of the animals (coral polyps) that build coral reefs. Coral reefs are highly diverse ecological communities. That diversity ultimately depends on corals—and on the mutualistic protists that nourish them. Corals support reef diversity by providing food to some species and habitat to many others.

Another example is the wood-digesting protists that inhabit the gut of many termite species **(Figure 28.27)**. Unaided, termites cannot digest wood, and they rely on protistan or prokaryotic symbionts to do so. Termites cause over $3.5 billion in damage annually to wooden homes in the United States.

Symbiotic protists also include parasites that have compromised the economies of entire countries. Consider the malaria-causing protist *Plasmodium*: Income levels in countries hard hit by malaria are 33% lower than in similar countries free of the disease. Protists can have devastating effects on other species too. Massive fish kills have been attributed to *Pfiesteria shumwayae* (see Figure 28.15), a dinoflagellate parasite that attaches to its victims and eats their skin. Among species that parasitize plants, the stramenopile *Phytophthora ramorum* has emerged as a major new forest pathogen. This species causes sudden oak death (SOD), a disease that has killed millions of oaks and other trees in the United States and Great Britain (**Figure 28.28**; also see Concept 54.5). A closely related species, *P. infestans*, causes potato late blight, which turns the stalks and stems of potato plants into black slime. Late blight contributed to the devastating Irish famine of the 19th century, in which a million people died and at least that many were forced to leave Ireland. The disease continues to be a major problem today, causing crop losses as high as 70% in some regions.



➤ Figure 28.27 A symbiotic protist. This organism is a hypermastigote, a member of a group of parabasalids that live in the gut of termites and certain cockroaches and enable the hosts to digest wood (SEM).

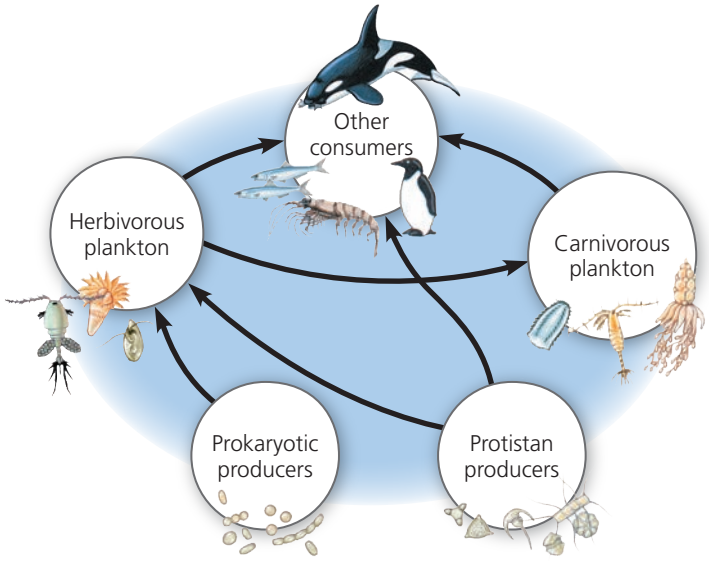
➤ Figure 28.28 Sudden oak death. Dead oak trees are visible in this landscape in Essex, England. Infected trees lose their ability to adjust to cycles of wet and dry weather.

Photosynthetic Protists

Many protists are important **producers**, organisms that use energy from light (or in some prokaryotes, inorganic chemicals) to convert CO_2 to organic compounds. Producers form the base of ecological food webs. In aquatic communities, the main producers are photosynthetic protists and prokaryotes **(Figure 28.29)**. All other organisms in the community depend on them for food, either directly (by eating them) or indirectly (by eating an organism that ate a producer). Scientists estimate that roughly 30% of the world's photosynthesis is performed by diatoms, dinoflagellates, multicellular algae, and other aquatic protists. Photosynthetic prokaryotes contribute another 20%, and plants are responsible for the remaining 50%.

Because producers form the foundation of food webs, factors that affect producers can dramatically affect their entire community. In aquatic environments, photosynthetic protists are often held in check by low concentrations of nitrogen, phosphorus, or iron. Various human actions can increase the

Figure 28.29 Protists: key producers in aquatic communities. Arrows in this simplified food web lead from food sources to the organisms that eat them.



concentrations of these elements in aquatic communities. For example, when fertilizer is applied to a field, some of the fertilizer may be washed by rainfall into a river that drains into a lake or ocean. When people add nutrients to aquatic communities in this or other ways, the abundance of photosynthetic protists can increase spectacularly. Such increases can have major ecological consequences, including the formation of large "dead zones" in marine ecosystems (see Figure 56.23).

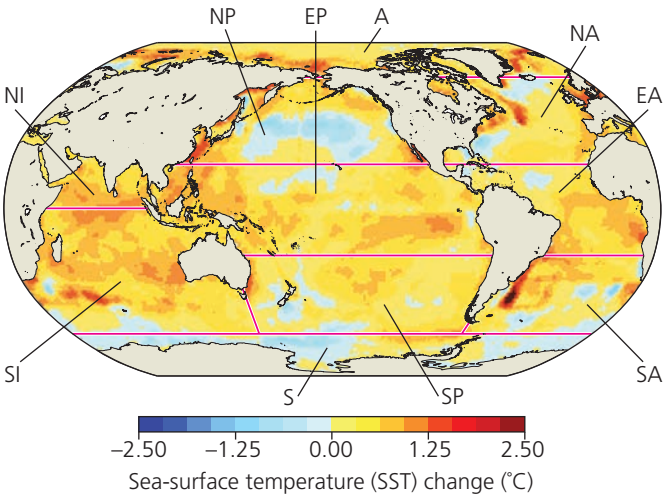
A pressing question is how global warming will affect photosynthetic protists and other producers. As shown in **Figure 28.30**, the growth and biomass of photosynthetic protists and prokaryotes have declined in many ocean regions as sea surface temperatures have increased. By what mechanism do rising sea surface temperatures reduce the growth of marine producers? One hypothesis relates to the rise or upwelling of cold, nutrient-rich waters from below. Many marine producers rely on nutrients brought to the surface in this way. However, rising sea surface temperatures can cause the formation of a layer of light, warm water that acts as a barrier to nutrient upwelling—thus reducing the growth of marine producers. If sustained, the changes shown in Figure 28.30 would likely have far-reaching effects on marine ecosystems, fishery yields, and the global carbon cycle (see Figure 55.14). Global warming can also affect producers on land, but there the base of food webs is occupied not by protists but by plants, which we will discuss in Chapters 29 and 30.

CONCEPT CHECK 28.6

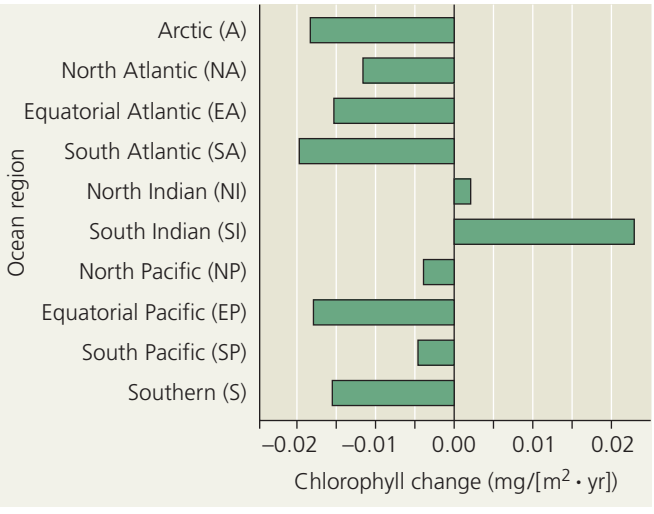
1. Justify the claim that photosynthetic protists are among the biosphere's most important organisms.
2. Describe three symbioses that include protists.
3. **WHAT IF? ➤** High water temperatures and pollution can cause corals to expel their dinoflagellate symbionts. How might such "coral bleaching" affect corals and other species?
4. **MAKE CONNECTIONS ➤** The bacterium *Wolbachia* is a symbiont that lives in mosquito cells and spreads rapidly through mosquito populations. *Wolbachia* can make mosquitoes resistant to infection by *Plasmodium*; researchers are seeking a strain that confers resistance and does not harm mosquitoes. Compare evolutionary changes that could occur if malaria control is attempted using such a *Wolbachia* strain versus using insecticides to kill mosquitoes. (Review Figure 28.16 and Concept 23.4.)

For suggested answers, see Appendix A.

Figure 28.30 Effects of climate change on marine producers.



(a) Researchers studied 10 ocean regions, identified with letters on the map (see (b) for the corresponding names). SSTs have increased since 1950 in most areas of these regions.



(b) The concentration of chlorophyll, an index for the biomass and growth of marine producers, has decreased over the same time period in most ocean regions.

28 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

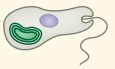
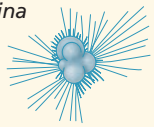
CONCEPT 28.1

Most eukaryotes are single-celled organisms *(pp. 646–651)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

- Domain Eukarya includes many groups of **protists**, along with plants, animals, and fungi. Unlike prokaryotes, protists and other eukaryotes have a nucleus and other membrane-enclosed organelles, as well as a cytoskeleton that enables them to have asymmetric forms and to change shape as they feed, move, or grow.
- Protists are structurally and functionally diverse and have a wide variety of life cycles. Most are unicellular. Protists include photoautotrophs, heterotrophs, and **mixotrophs**.
- Current evidence indicates that eukaryotes originated by **endosymbiosis** when an archaeal host (or a host closely related to the archaeans) engulfed an alpha proteobacterium that would evolve into an organelle found in all eukaryotes, the mitochondrion.
- Plastids are thought to be descendants of cyanobacteria that were engulfed by early eukaryotic cells. The plastid-bearing lineage eventually evolved into red algae and green algae. Other protist groups evolved from secondary endosymbiotic events in which red algae or green algae were themselves engulfed.
- In one hypothesis, eukaryotes are grouped into four supergroups, each a monophyletic clade: Excavata, SAR, Archaeplastida, and Unikonta.

? *Describe similarities and differences between protists and other eukaryotes.*

Key Concept/Eukaryote Supergroup	Major Groups	Key Morphological Characteristics	Specific Examples
CONCEPT 28.2 **Excavates include protists with modified mitochondria and protists with unique flagella** *(pp. 651–653)* ? *What evidence indicates that the excavates form a clade?*	**Diplomonads and parabasalids**	Modified mitochondria	*Giardia, Trichomonas* 
	Euglenozoans Kinetoplastids Euglenids	Spiral or crystalline rod inside flagella	*Trypanosoma, Euglena* 
CONCEPT 28.3 **SAR is a highly diverse group of protists defined by DNA similarities** *(pp. 653–660)* ? *Although they are not photosynthetic, apicomplexan parasites such as* Plasmodium *have modified plastids. Describe a current hypothesis that explains this observation.*	**Stramenopiles** Diatoms Golden algae Brown algae	Hairy and smooth flagella	*Phytophthora, Laminaria* 
	Alveolates Dinoflagellates Apicomplexans Ciliates	Membrane-enclosed sacs (alveoli) beneath plasma membrane	*Pfiesteria, Plasmodium, Paramecium* 
	Rhizarians Radiolarians Forams Cercozoans	Amoebas with threadlike pseudopodia	*Globigerina* 
CONCEPT 28.4 **Red algae and green algae are the closest relatives of plants** *(pp. 660–662)* ? *On what basis do systematists place plants in the same supergroup (Archaeplastida) as red and green algae?*	**Red algae**	Phycoerythrin (photosynthetic pigment)	*Porphyra* 
	Green algae	Plant-type chloroplasts	*Chlamydomonas, Ulva* 
	Plants	(See Chapters 29 and 30.)	Mosses, ferns, conifers, flowering plants 
CONCEPT 28.5 **Unikonts include protists that are closely related to fungi and animals** *(pp. 662–666)* ? *Describe a key feature for each of the main protist subgroups of Unikonta.*	**Amoebozoans** Slime molds Tubulinids Entamoebas	Amoebas with lobe-shaped or tube-shaped pseudopodia	*Amoeba, Dictyostelium* 
	Opisthokonts	(Highly variable; see Chapters 31–34.)	Choanoflagellates, nucleariids, animals, fungi 

CONCEPT 28.6

Protists play key roles in ecological communities *(pp. 666–667)*

- Protists form a wide range of mutualistic and parasitic relationships that affect their symbiotic partners and many other members of the community.
- Photosynthetic protists are among the most important **producers** in aquatic communities. Because they are at the base of the food web, factors that affect photosynthetic protists affect many other species in the community.

? *Describe several protists that are ecologically important.*

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–6 can be found in the Study Area in MasteringBiology.

7. EVOLUTION CONNECTION • DRAW IT Medical researchers seek to develop drugs that can kill or restrict the growth of human pathogens yet have few harmful effects on patients. These drugs often work by disrupting the metabolism of the pathogen or by targeting its structural features.

PRACTICE TEST
goo.gl/iAsVgL

Draw and label a phylogenetic tree that includes an ancestral prokaryote and the following groups of organisms: Excavata, SAR, Archaeplastida, Unikonta, and, within Unikonta, amoebozoans, animals, choanoflagellates, fungi, and nucleariids. Based on this tree, hypothesize whether it would be most difficult to develop drugs to combat human pathogens that are prokaryotes, protists, animals, or fungi. (You do not need to consider the evolution of drug resistance by the pathogen.)

8. SCIENTIFIC INQUIRY Applying the "If…then" logic of deductive reasoning (see Concept 1.3), what are a few of the predictions that arise from the hypothesis that plants evolved from green algae? Put another way, how could you test this hypothesis?

9. WRITE ABOUT A THEME: INTERACTIONS Organisms interact with each other and the physical environment. In a short essay (100–150 words), explain how the response of diatom populations to a drop in nutrient availability can affect both other organisms and aspects of the physical environment (such as carbon dioxide concentrations).

10. SYNTHESIZE YOUR KNOWLEDGE

This micrograph shows a single-celled eukaryote, the ciliate *Didinium* (left), about to engulf its *Paramecium* prey, which is also a ciliate. Identify the eukaryotic supergroup to which ciliates belong and describe the role of endosymbiosis in the evolutionary history of that supergroup. Are these ciliates more closely related to all other protists than they are to plants, fungi, or animals? Explain.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

Nonvascular and Seedless Vascular Plants

29

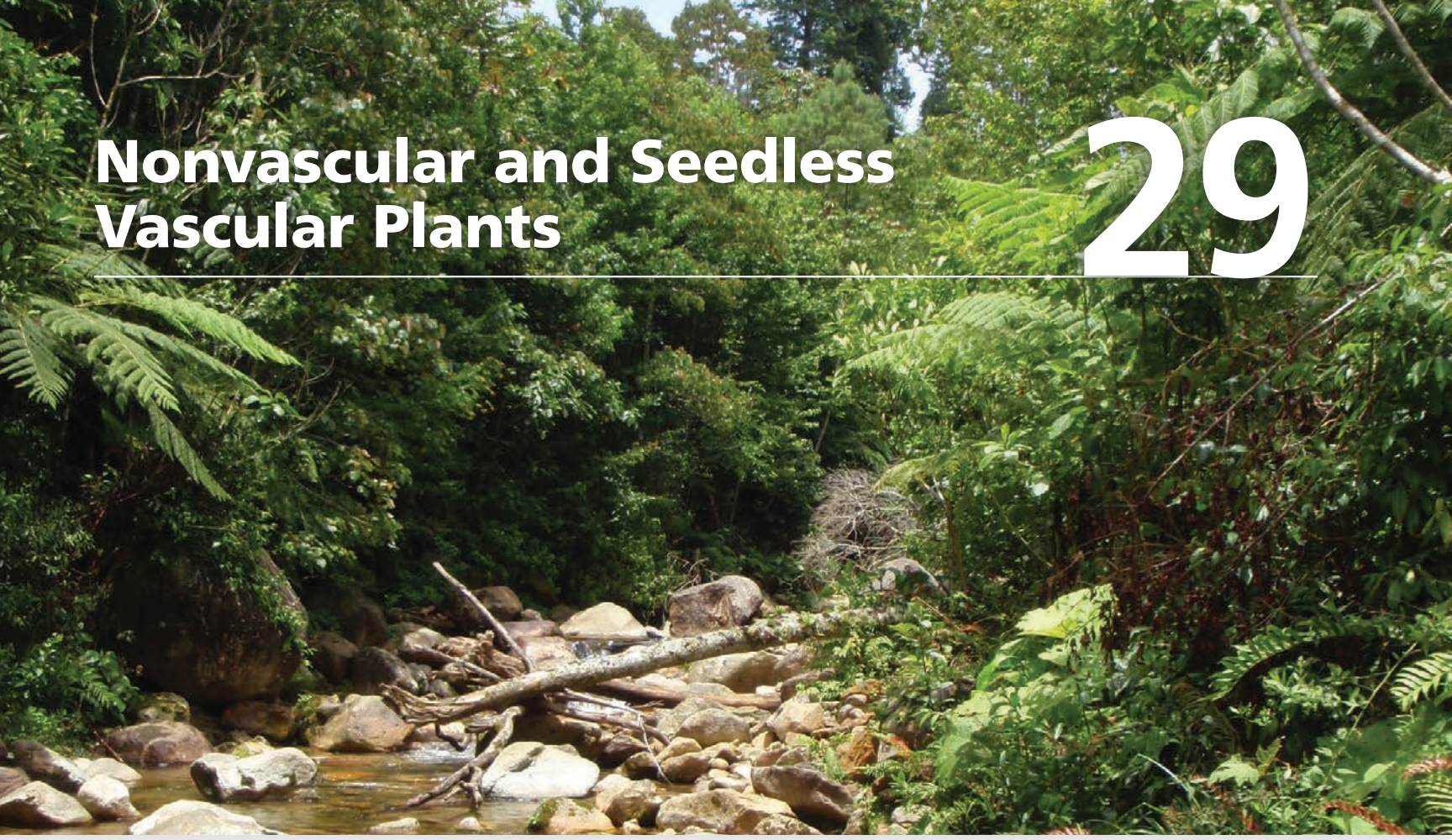


Figure 29.1 How did plants change the world?

KEY CONCEPTS

29.1 Plants evolved from green algae

29.2 Mosses and other nonvascular plants have life cycles dominated by gametophytes

29.3 Ferns and other seedless vascular plants were the first plants to grow tall

The Greening of Earth

Looking at a lush landscape, such as that shown in **Figure 29.1**, it is hard to imagine the land without plants or other organisms. Yet for much of Earth's history, the land was largely lifeless. Geochemical analysis and fossil evidence suggest that thin coatings of cyanobacteria and protists existed on land by 1.2 billion years ago. But it was only within the last 500 million years that small plants, fungi, and animals joined them ashore. Finally, by about 385 million years ago, tall plants appeared, leading to the first forests.

Today, there are more than 290,000 known plant species. Plants inhabit all but the harshest environments, such as some mountaintop and desert areas and the polar ice sheets. Although a few plant species, such as sea grasses, returned to aquatic habitats during their evolution, most present-day plants live on land. In this text, we distinguish plants from algae, which are photosynthetic protists.

Plants enabled other life-forms to survive on land. For example, plants supply oxygen and are a key source of food for terrestrial animals. Also, by their very presence, plants such as the trees of a forest physically create the habitats required by animals and many other organisms. This chapter traces the first 100 million years of plant evolution, including the emergence of seedless plants such as mosses and ferns. Chapter 30 examines the later evolution of seed plants.

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

CONCEPT 29.1

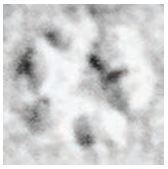
Plants evolved from green algae

As you read in Chapter 28, green algae called charophytes are the closest relatives of plants. We'll begin with a closer look at the evidence for this relationship.

Morphological and Molecular Evidence

Many key traits of plants also appear in some algae. For example, plants are multicellular, eukaryotic, photosynthetic autotrophs, as are brown, red, and certain green algae. Plants have cell walls made of cellulose, and so do green algae, dinoflagellates, and brown algae. And chloroplasts with chlorophylls *a* and *b* are present in green algae, euglenids, and a few dinoflagellates, as well as in plants.

However, the charophytes are the only present-day algae that share the following distinctive traits with plants, suggesting that they are the closest living relatives of plants:



- **Rings of cellulose-synthesizing proteins.** The cells of both plants and charophytes have distinctive circular rings of proteins (small photo) embedded in the plasma membrane. These protein rings synthesize the cellulose microfibrils of the cell wall. In contrast, noncharophyte algae have linear sets of proteins that synthesize cellulose.
- **Structure of flagellated sperm.** In species of plants that have flagellated sperm, the structure of the sperm closely resembles that of charophyte sperm.
- **Formation of a phragmoplast.** Particular details of cell division occur only in plants and certain charophytes. For example, a group of microtubules known as the phragmoplast forms between the daughter nuclei of a dividing cell. A cell plate then develops in the middle of the phragmoplast, across the midline of the dividing cell (see Figure 12.10). The cell plate, in turn, gives rise to a new cross wall that separates the daughter cells.

Studies of nuclear, chloroplast, and mitochondrial DNA from a wide range of plants and algae indicate that certain groups of charophytes—such as *Zygnema* (see Figure 28.22a) and *Coleochaete*—are the closest living relatives of plants. Although this evidence shows that plants arose from within a group of charophyte algae, it does not mean that plants are descended from these living algae. Even so, present-day charophytes may tell us something about the algal ancestors of plants.

Adaptations Enabling the Move to Land

Many species of charophyte algae inhabit shallow waters around the edges of ponds and lakes, where they are subject to occasional drying. In such environments, natural selection favors individual algae that can survive periods when they are not submerged. In charophytes, a layer of a durable polymer called **sporopollenin** prevents exposed zygotes from drying out. A similar chemical adaptation is found in the tough sporopollenin walls that encase plant spores.

The accumulation of such traits by at least one population of charophyte algae (now extinct) probably enabled their descendants—the first plants—to live permanently above the waterline. This ability opened a new frontier: a terrestrial habitat that offered enormous benefits. The bright sunlight was unfiltered by water and plankton; the atmosphere offered more plentiful carbon dioxide than did water; and the soil by the water's edge was rich in some mineral nutrients. But these benefits were accompanied by challenges: a relative scarcity of water and a lack of structural support against gravity. (To appreciate why such support is important, picture how the soft body of a jellyfish sags when taken out of water.) Plants diversified as new adaptations arose that enabled them to thrive despite these challenges.

Today, what adaptations are unique to plants? The answer depends on where you draw the boundary dividing plants from algae **(Figure 29.2)**. Since the placement of this boundary is the subject of ongoing debate, this text uses a traditional definition that equates the kingdom Plantae with embryophytes (plants with embryos). In this context, let's now examine the derived traits that separate plants from their closest algal relatives.

Derived Traits of Plants

Several adaptations that facilitate survival and reproduction on dry land emerged after plants diverged from their algal relatives. **Figure 29.3** depicts five such traits that are found in plants but not in charophyte algae.

Figure 29.2 Three possible "plant" kingdoms.

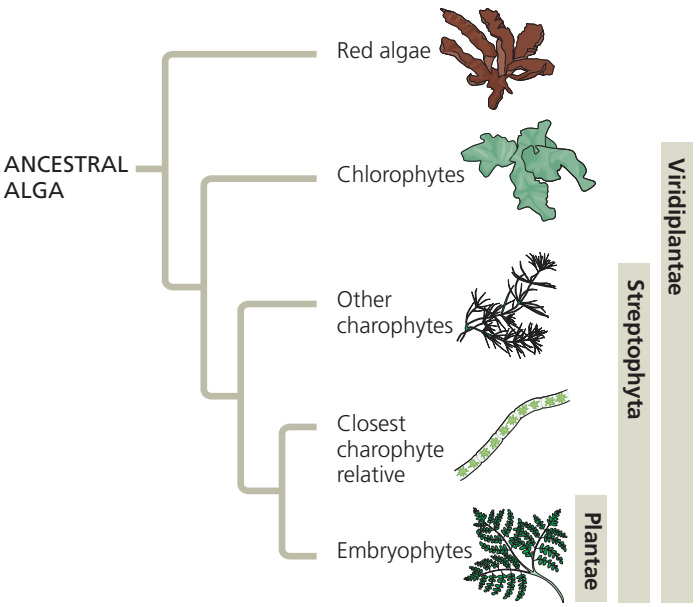


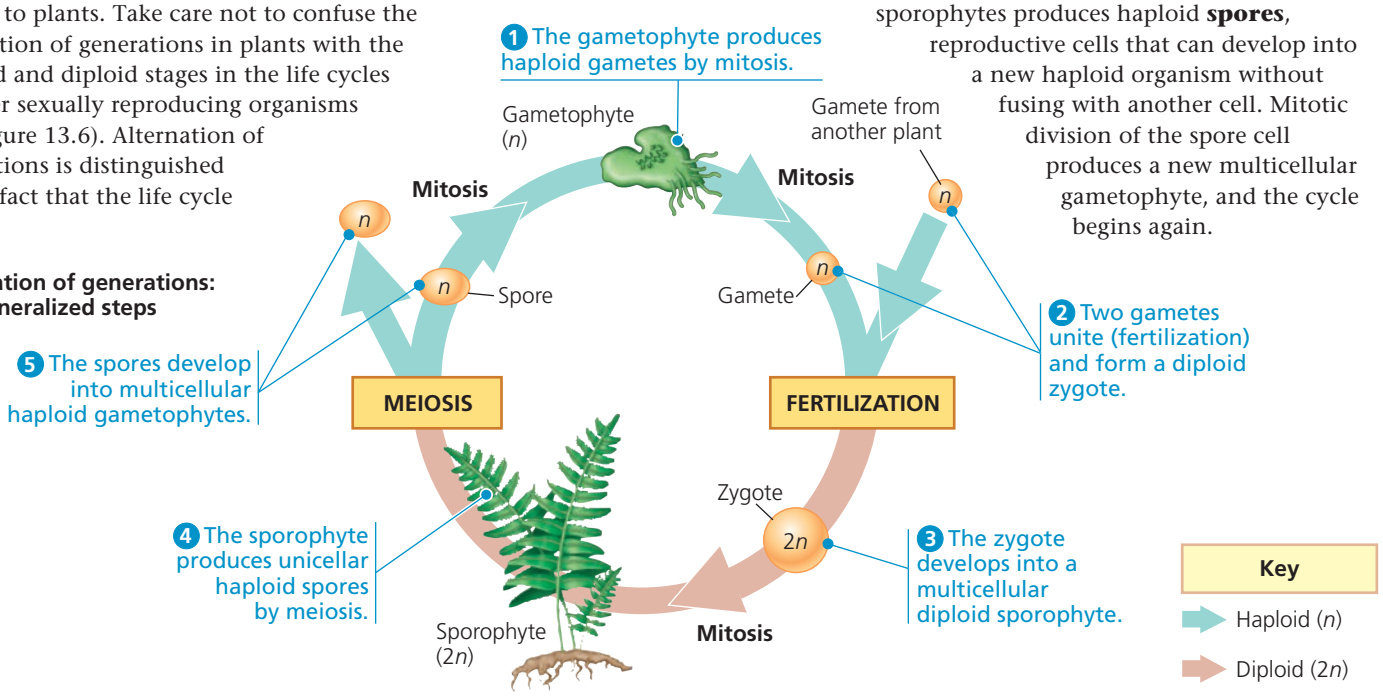
Figure 29.3 Exploring Derived Traits of Plants

Charophyte algae lack the key traits of plants described in this figure: alternation of generations; multicellular, dependent embryos; walled spores produced in sporangia; multicellular gametangia; and apical meristems. This suggests that these traits were absent in the ancestor common to plants and charophytes but instead evolved as derived traits of plants. Not every plant exhibits all of these traits; certain lineages of plants have lost some traits over time.

Alternation of Generations

The life cycles of all plants alternate between two generations of distinct multicellular organisms: gametophytes and sporophytes. As shown in the diagram below (using a fern as an example), each generation gives rise to the other, a process that is called **alternation of generations**. This type of reproductive cycle evolved in various groups of algae but does not occur in the charophytes, the algae most closely related to plants. Take care not to confuse the alternation of generations in plants with the haploid and diploid stages in the life cycles of other sexually reproducing organisms (see Figure 13.6). Alternation of generations is distinguished by the fact that the life cycle includes both multicellular haploid organisms and multicellular diploid organisms. The multicellular haploid **gametophyte** ("gamete-producing plant") is named for its production by mitosis of haploid gametes—eggs and sperm—that fuse during fertilization, forming diploid zygotes. Mitotic division of the zygote produces a multicellular diploid **sporophyte** ("spore-producing plant"). Meiosis in a mature sporophytes produces haploid **spores**, reproductive cells that can develop into a new haploid organism without fusing with another cell. Mitotic division of the spore cell produces a new multicellular gametophyte, and the cycle begins again.

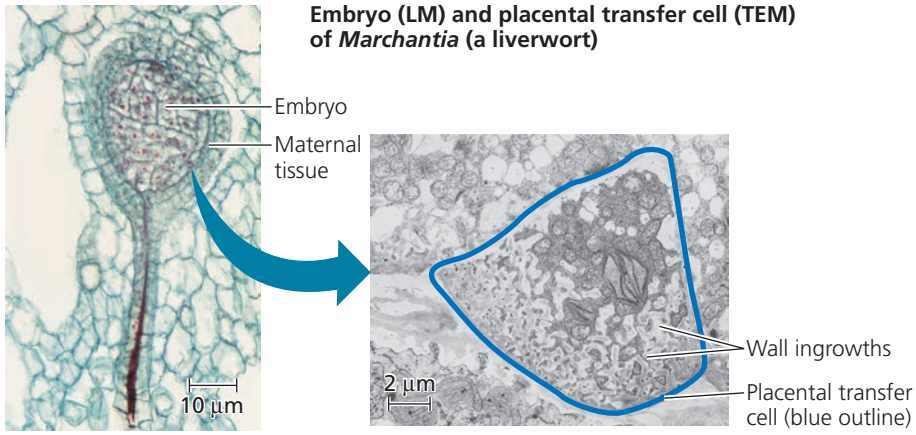
Alternation of generations: five generalized steps



Multicellular, Dependent Embryos

As part of a life cycle with alternation of generations, multicellular plant embryos develop from zygotes that are retained within the tissues of the female parent (a gametophyte). The parental tissues protect the developing embryo from harsh environmental conditions and provide nutrients such as sugars and amino acids. The embryo has specialized *placental transfer cells* that enhance the transfer of nutrients to the embryo through elaborate ingrowths of the wall surface (plasma membrane and cell wall). The multicellular, dependent embryo of plants is such a significant derived trait that plants are also known as **embryophytes**.

Embryo (LM) and placental transfer cell (TEM) of *Marchantia* (a liverwort)

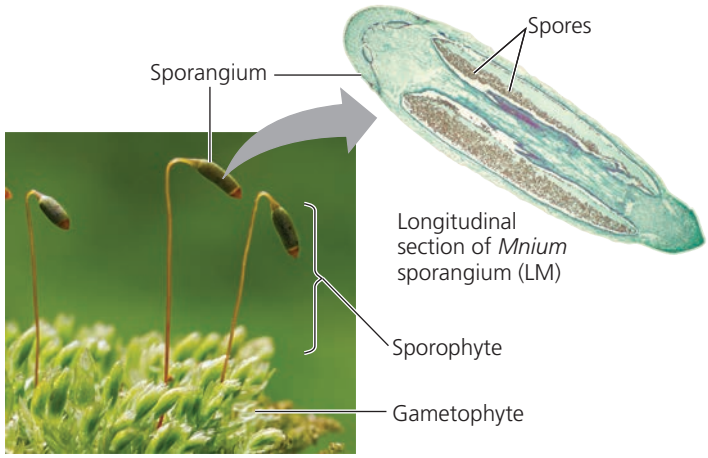


MAKE CONNECTIONS ➤ *Review sexual life cycles in Figure 13.6. Identify which type of sexual life cycle has alternation of generations, and summarize how it differs from other life cycles.*

Walled Spores Produced in Sporangia

Plant spores are haploid reproductive cells that can grow into multicellular haploid gametophytes by mitosis. The polymer sporopollenin makes the walls of plant spores tough and resistant to harsh environments. This chemical adaptation enables spores to be dispersed through dry air without harm.

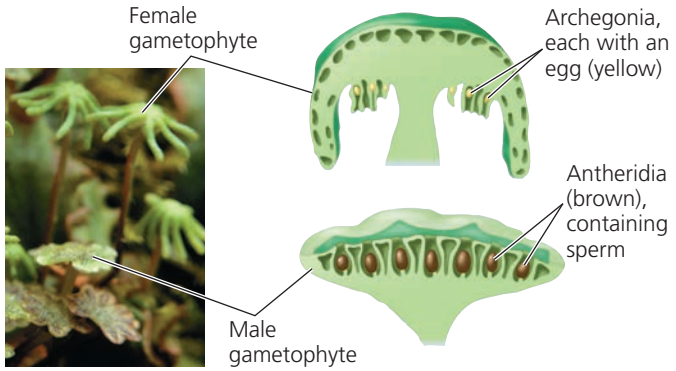
The sporophyte has multicellular organs called **sporangia** (singular, *sporangium*) that produce the spores. Within a sporangium, diploid cells called **sporocytes**, or spore mother cells, undergo meiosis and generate the haploid spores. The outer tissues of the sporangium protect the developing spores until they are released into the air. Multicellular sporangia that produce spores with sporopollenin-enriched walls are key terrestrial adaptations of plants. Although charophytes also produce spores, these algae lack multicellular sporangia, and their flagellated, water-dispersed spores lack sporopollenin.



Sporophytes and sporangia of *Mnium* (a moss)

Multicellular Gametangia

Another feature distinguishing early plants from their algal ancestors was the production of gametes within multicellular organs called **gametangia**. The female gametangia are called **archegonia** (singular, archegonium). Each archegonium is a pear-shaped organ that produces a single nonmotile egg retained within the bulbous part of the organ (the top for the species shown here). The male gametangia, called **antheridia** (singular, antheridium), produce sperm and release them into the environment. In many groups of present-day plants, the sperm have flagella and swim to the eggs through water droplets or a film of water. Each egg is fertilized within an archegonium, where the zygote develops into an embryo. The gametophytes of seed plants are so reduced in size (as you will see in Chapter 30) that the archegonia and antheridia have been lost in many lineages.

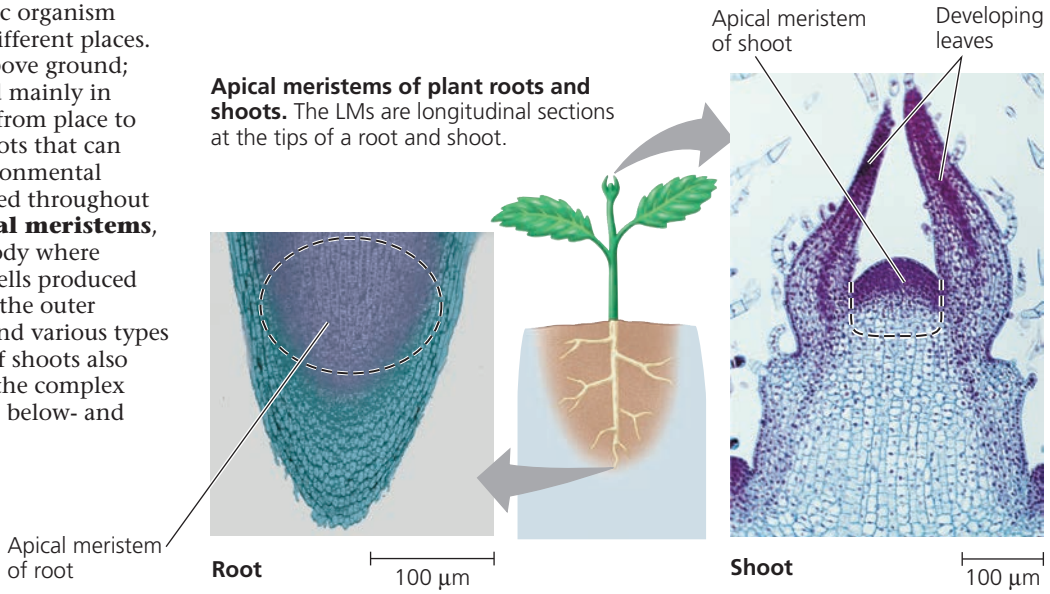


Archegonia and antheridia of *Marchantia* (a liverwort)

Apical Meristems

In terrestrial habitats, a photosynthetic organism finds essential resources in two very different places. Light and CO_2 are mainly available above ground; water and mineral nutrients are found mainly in the soil. Though plants cannot move from place to place, most plants have roots and shoots that can elongate, increasing exposure to environmental resources. Growth in length is sustained throughout the plant's life by the activity of **apical meristems**, regions at growing tips of the plant body where one or more cells divide repeatedly. Cells produced by apical meristems differentiate into the outer epidermis, which protects the body, and various types of internal tissues. Apical meristems of shoots also generate leaves in most plants. Thus, the complex bodies of most plants have specialized below- and aboveground organs.

Apical meristems of plant roots and shoots. The LMs are longitudinal sections at the tips of a root and shoot.



Additional derived traits that relate to terrestrial life have evolved in many plant species. For example, the epidermis in many species has a covering, the **cuticle**, that consists of wax and other polymers. Permanently exposed to the air, plants run a far greater risk of desiccation (drying out) than do their algal relatives. The cuticle acts as waterproofing, helping prevent excessive water loss from the aboveground plant organs, while also providing some protection from microbial attack. Most plants also have specialized pores called **stomata** (singular, *stoma*), which support photosynthesis by allowing the exchange of CO_2 and O_2 between the outside air and the plant (see Figure 11.4). Stomata are also the main avenues by which water evaporates from the plant; in hot, dry conditions, the stomata close, minimizing water loss.

The earliest plants lacked true roots and leaves. Without roots, how did these plants absorb nutrients from the soil? Fossils dating from 420 million years ago reveal an adaptation that may have aided early plants in nutrient uptake: They formed symbiotic associations with fungi. We'll describe these associations, called *mycorrhizae*, and their benefits to both plants and fungi in more detail in Concept 31.1. For now, the main point is that mycorrhizal fungi form extensive networks of filaments through the soil and transfer nutrients to their symbiotic plant partner. This benefit may have helped plants without roots to colonize land.

The Origin and Diversification of Plants

The algae most closely related to plants include many unicellular species and small colonial species. Since it is likely that the first plants were similarly small, the search for the earliest fossils of plants has focused on the microscopic world. As mentioned earlier, microorganisms colonized land as early as 1.2 billion years ago. But the microscopic fossils that document life on land changed dramatically 470 million years ago with the appearance of spores from early plants.

What distinguishes these spores from those of algae or fungi? One clue comes from their chemical composition, which matches the composition of plant spores today but differs from that of the spores of other organisms. In addition, the walls of these ancient spores have structural features that today are found only in the spores of certain plants (liverworts). And in rocks dating to 450 million years ago, researchers have discovered similar spores embedded in plant cuticle material that resembles spore-bearing tissue in living plants **(Figure 29.4)**.

Fossils of larger plant structures, such as the *Cooksonia* sporangium in **Figure 29.5**, date to 425 million years ago, which is 45 million years after the appearance of plant spores in the fossil record. While the precise age (and form) of the first plants has yet to be discovered, those ancestral species gave rise to the vast diversity of living plants. **Table 29.1** summarizes the ten extant phyla in the taxonomic scheme used in this text. (Extant lineages are those that have surviving members.) As you read the rest of this section, look at Table 29.1 together with **Figure 29.6**, which reflects a view of plant phylogeny that is based on plant morphology, biochemistry, and genetics.

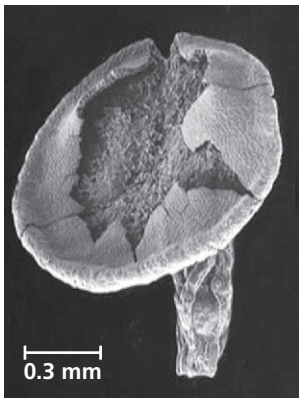
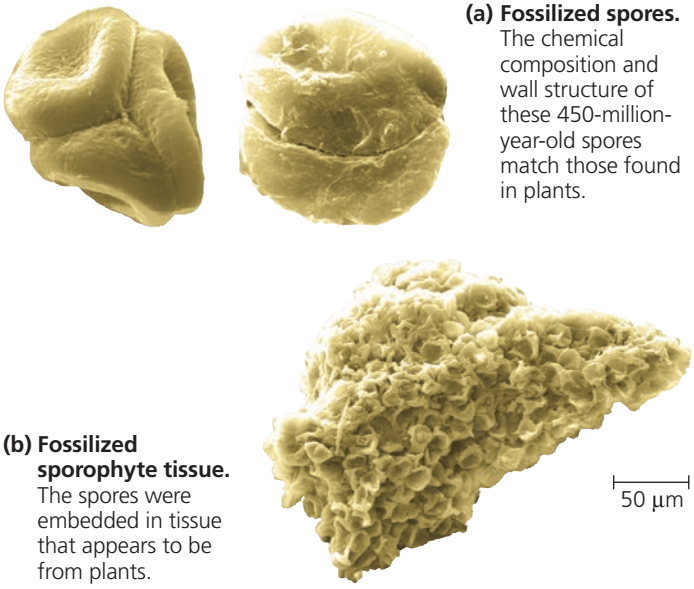


Figure 29.5 ***Cooksonia*** **sporangium fossil**

Figure 29.4 **Ancient plant spores and tissue** (colorized SEMs).



(a) Fossilized spores. The chemical composition and wall structure of these 450-million-year-old spores match those found in plants.

(b) Fossilized sporophyte tissue. The spores were embedded in tissue that appears to be from plants.

One way to distinguish groups of plants is whether or not they have an extensive system of **vascular tissue**, cells joined into tubes that transport water and nutrients throughout the plant body. Most present-day plants have a complex vascular tissue system and are therefore called **vascular plants**. Plants that do not have an extensive transport system—liverworts, mosses, and hornworts—are described as "nonvascular" plants, even though some mosses do have simple vascular tissue. Nonvascular plants are often informally called **bryophytes** (from the Greek *bryon*, moss, and *phyton*, plant). Although the term *bryophyte* is commonly used to refer to all nonvascular plants, molecular studies and morphological analyses of sperm structure have concluded that bryophytes do not form a monophyletic group (a clade).

Vascular plants, which form a clade that comprises about 93% of all extant plant species, can be categorized further into smaller clades. Two of these clades are the **lycophytes** (the club mosses and their relatives) and the **monilophytes**

Table 29.1 Ten Phyla of Extant Plants

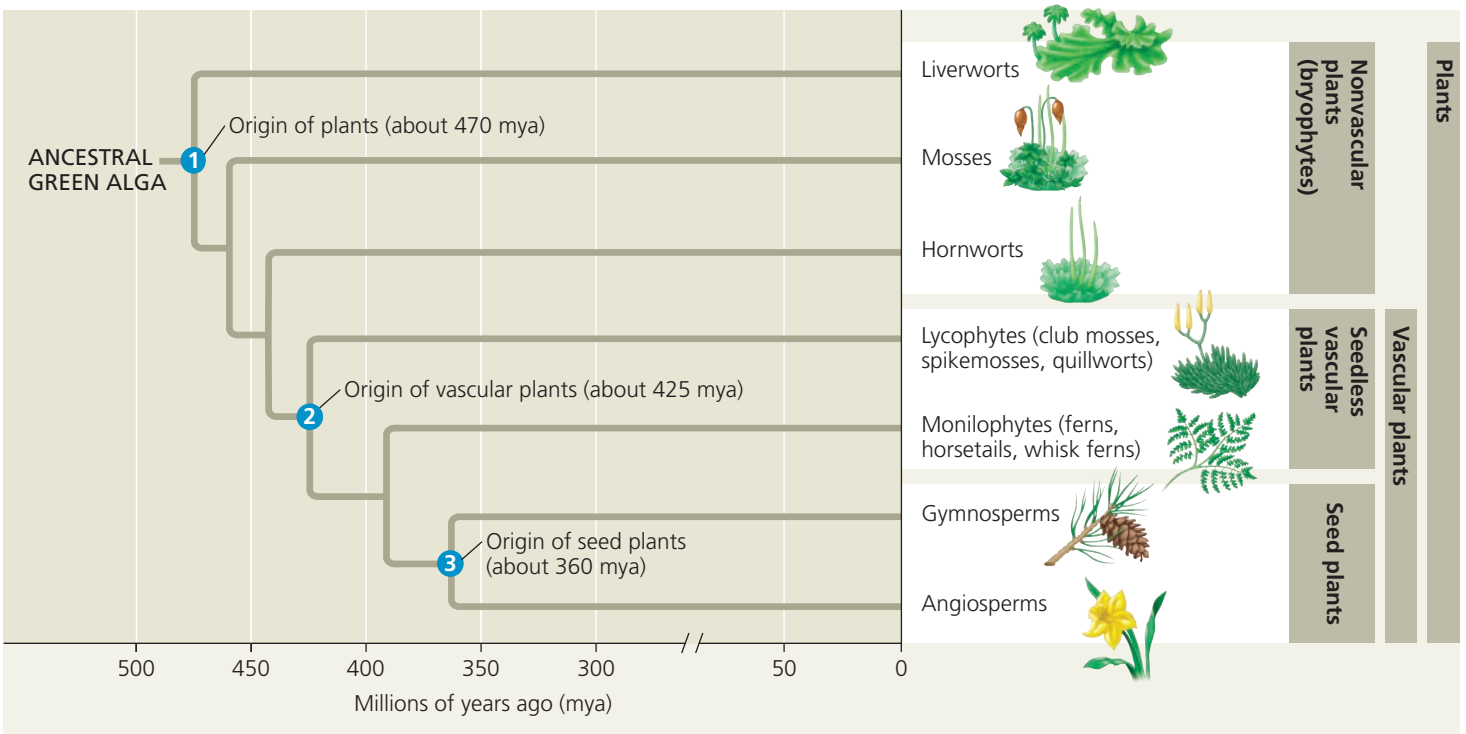
	Common Name	Number of Known Species
Nonvascular Plants (Bryophytes)		
Phylum Hepatophyta	Liverworts	9,000
Phylum Bryophyta	Mosses	15,000
Phylum Anthocerophyta	Hornworts	100
Vascular Plants		
Seedless Vascular Plants		
Phylum Lycophyta	Lycophytes	1,200
Phylum Monilophyta	Monilophytes	12,000
Seed Plants		
Gymnosperms		
Phylum Ginkgophyta	Ginkgo	1
Phylum Cycadophyta	Cycads	130
Phylum Gnetophyta	Gnetophytes	75
Phylum Coniferophyta	Conifers	600
Angiosperms		
Phylum Anthophyta	Flowering plants	250,000

(ferns and their relatives). The plants in each of these clades lack seeds, which is why collectively the two clades are often informally called **seedless vascular plants**. However, notice in Figure 29.6 that, like bryophytes, seedless vascular plants do not form a clade.

A group such as the bryophytes or the seedless vascular plants is sometimes referred to as a *grade*, a collection of organisms that share key biological features. Grades can be informative by grouping organisms according to their features, such as having a vascular system but lacking seeds. But members of a grade, unlike members of a clade, do not necessarily share the same ancestry. For example, even though monilophytes and lycophytes are all seedless vascular plants, monilophytes share a more recent common ancestor with seed plants. As a result, we would expect monilophytes and seed plants to share key traits not found in lycophytes—and they do, as you'll read in Concept 29.3.

A third clade of vascular plants consists of seed plants, which represent the vast majority of living plant species. A **seed** is an embryo packaged with a supply of nutrients inside a protective coat. Seed plants can be divided into two groups, gymnosperms and angiosperms, based on the absence or presence of enclosed chambers in which seeds mature. **Gymnosperms** (from the Greek *gymnos*, naked, and *sperm*, seed) are known as "naked seed" plants because their seeds are not enclosed in chambers. Living gymnosperm species, the most familiar of which are the conifers, form a clade. **Angiosperms** (from the Greek *angion*, container) are a huge clade consisting of all flowering plants; their seeds develop

Figure 29.6 Highlights of plant evolution. The phylogeny shown here illustrates a leading hypothesis about the relationships between plant groups.



MAKE CONNECTIONS ➤ *The figure identifies which lineages are plants, nonvascular plants, vascular plants, seedless vascular plants, and seed plants. Which of these categories are monophyletic, and which are paraphyletic? Explain. (See Figure 22.10 to review these terms.)*

inside chambers that originate within flowers. Nearly 90% of living plant species are angiosperms.

Note that the phylogeny depicted in Figure 29.6 focuses only on the relationships between extant plant lineages. Paleobotanists have also discovered fossils belonging to extinct plant lineages. As you'll read later in the chapter, these fossils can reveal intermediate steps in the emergence of plant groups found on Earth today.

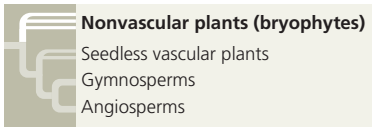
CONCEPT CHECK 29.1

1. Why do researchers identify the charophytes rather than another group of algae as the closest living relatives of plants?
2. Identify four derived traits that distinguish plants from charophyte green algae *and* facilitate life on land. Explain.
3. **WHAT IF?** ➤ What would the human life cycle be like if we had alternation of generations? Assume that the multicellular diploid stage would be similar in form to an adult human.

For suggested answers, see Appendix A.

CONCEPT 29.2

Mosses and other nonvascular plants have life cycles dominated by gametophytes



The nonvascular plants (bryophytes) are represented today by three phyla of small, herbaceous (nonwoody) plants: **liverworts** (phylum Hepatophyta), **mosses** (phylum Bryophyta), and **hornworts** (phylum Anthocerophyta). Liverworts and hornworts are named for their shapes, plus the suffix *wort* (from the Anglo-Saxon for "herb"). Mosses are familiar to many people, although some plants commonly called "mosses" are not really mosses at all. These include Irish moss (a red seaweed), reindeer moss (a lichen), club mosses (seedless vascular plants), and Spanish mosses (lichens in some regions and flowering plants in others).

Phylogenetic analyses indicate that liverworts, mosses, and hornworts diverged from other plant lineages early in the history of plant evolution (see Figure 29.6). Fossil evidence provides some support for this idea: The earliest spores of plants (dating from 450 to 470 million years ago) have structural features found only in the spores of liverworts, and by 430 million years ago spores similar to those of mosses and hornworts also occur in the fossil record. The earliest fossils of vascular plants date to about 425 million years ago.

Over the long course of their evolution, liverworts, mosses, and hornworts have acquired many unique adaptations. Next, we'll examine some of those features.

Bryophyte Gametophytes

Unlike vascular plants, in all three bryophyte phyla the haploid gametophytes are the dominant stage of the life cycle: They are usually larger and longer-living than the sporophytes, as shown in the moss life cycle in **Figure 29.7**. The sporophytes are typically present only part of the time.

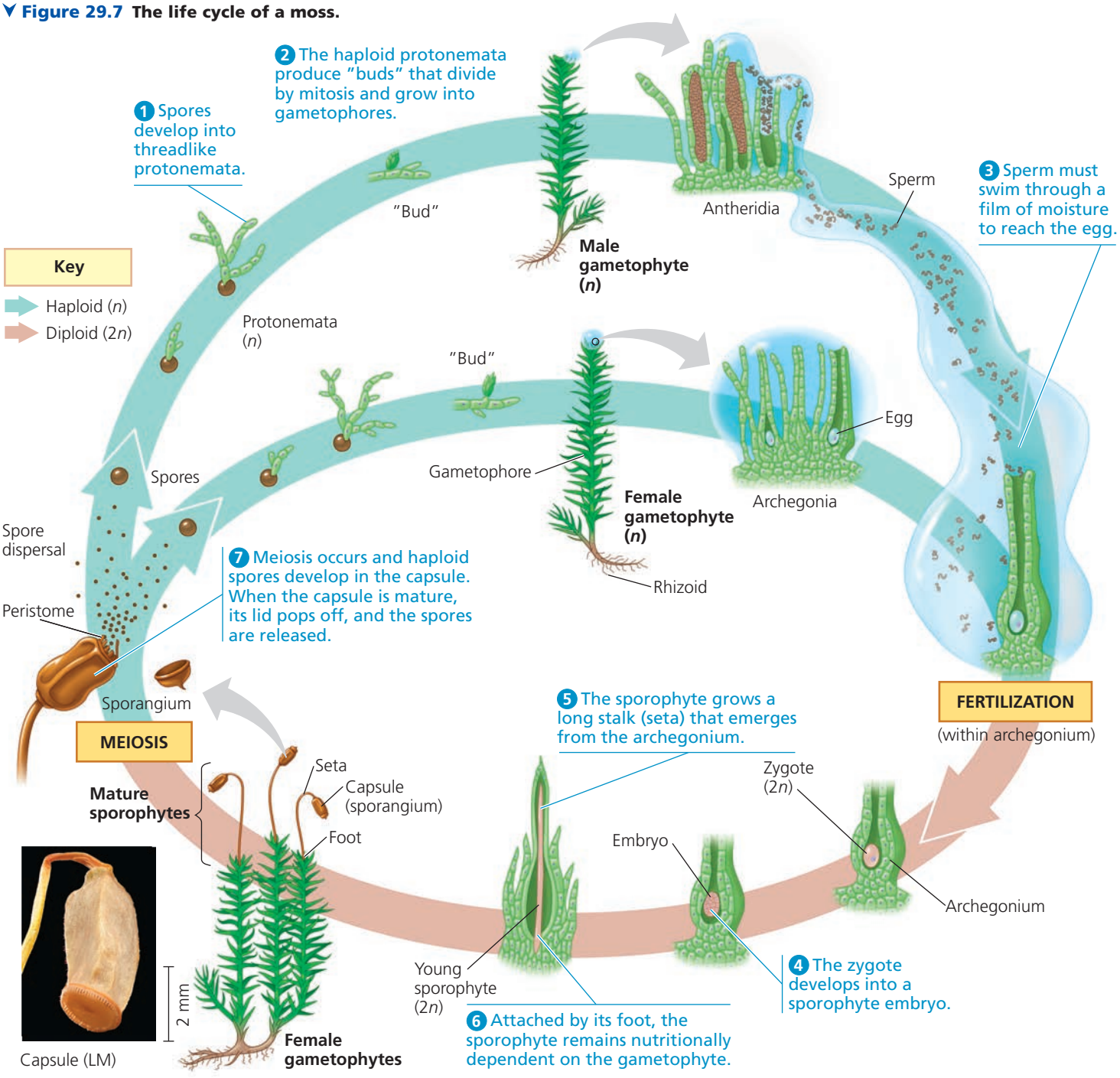
When bryophyte spores are dispersed to a favorable habitat, such as moist soil or tree bark, they may germinate and grow into gametophytes. Germinating moss spores, for example, characteristically produce a mass of green, branched, one-cell-thick filaments known as a **protonema** (plural, *protonemata*). A protonema has a large surface area that enhances absorption of water and minerals. In favorable conditions, a protonema produces one or more "buds." (Note that when referring to nonvascular plants, we often use quotation marks for structures similar to the buds, stems, and leaves of vascular plants because the definitions of these terms are based on vascular plant organs.) Each of these bud-like growths has an apical meristem that generates a gamete-producing structure known as a **gametophore**. Together, a protonema and one or more gametophores make up the body of a moss gametophyte.

Bryophyte gametophytes generally form ground-hugging carpets, partly because their body parts are too thin to support a tall plant. A second constraint on the height of many bryophytes is the absence of vascular tissue, which would enable long-distance transport of water and nutrients. (The thin structure of bryophyte organs makes it possible to distribute materials for short distances without specialized vascular tissue.) However, some mosses have conducting tissues in the center of their "stems." A few of these mosses can grow as tall as 60 cm (2 feet) as a result. Phylogenetic analyses suggest that conducting tissues similar to those of vascular plants arose independently in these mosses by convergent evolution.

The gametophytes are anchored by delicate **rhizoids**, which are long, tubular single cells (in liverworts and hornworts) or filaments of cells (in mosses). Unlike roots, which are found in vascular plant sporophytes, rhizoids are not composed of tissues. Bryophyte rhizoids also lack specialized conducting cells and do not play a primary role in water and mineral absorption.

Gametophytes can form multiple gametangia, each of which produces gametes and is covered by protective tissue. Each archegonium produces one egg, whereas each antheridium produces many sperm. Some bryophyte gametophytes are bisexual, but in mosses the archegonia and antheridia are

Figure 29.7 The life cycle of a moss.



VISUAL SKILLS ➤ *In this diagram, does the sperm cell that fertilizes the egg cell differ genetically from the egg? Explain.*

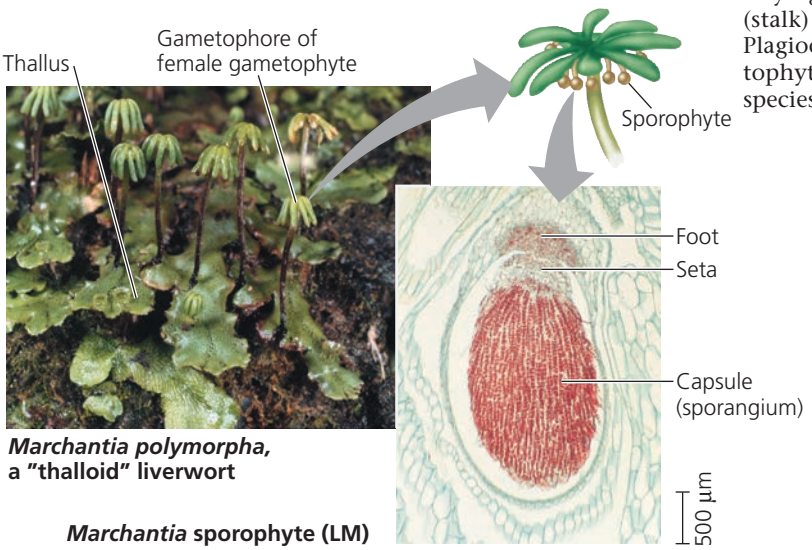
Animation: Moss Life Cycle

typically carried on separate female and male gametophytes. Flagellated sperm swim through a film of water toward eggs, entering the archegonia in response to chemical attractants. Eggs are not released but instead remain within the bases of archegonia. After fertilization, embryos are retained within the archegonia. Layers of placental transfer cells help transport nutrients to the embryos as they develop into sporophytes.

Bryophyte sperm typically require a film of water to reach the eggs. Given this requirement, it is not surprising that many bryophyte species are found in moist habitats. The fact that sperm swim through water to reach the egg also means that in species with separate male and female gametophytes (most species of mosses), sexual reproduction is likely to be more successful when individuals are located close to one another.

Liverworts (Phylum Hepatophyta)

This phylum's common and scientific names (from the Latin hepaticus, liver) refer to the liver-shaped gametophytes of its members, such as Marchantia, shown below. In medieval times, their shape was thought to be a sign that the plants could help treat liver diseases. Some liverworts, including *Marchantia*, are described as "thalloid" because of the flattened shape of their gametophytes. Marchantia gametangia are elevated on gametophores that look like miniature trees. You would need a magnifying glass to see the sporophytes, which have a short seta (stalk) with an oval or round capsule. Other liverworts, such as Plagiochila are called "leafy" because their stemlike gametophytes have many leaflike appendages. There are many more species of leafy liverworts than thalloid liverworts.



***Marchantia polymorpha*, a "thalloid" liverwort**

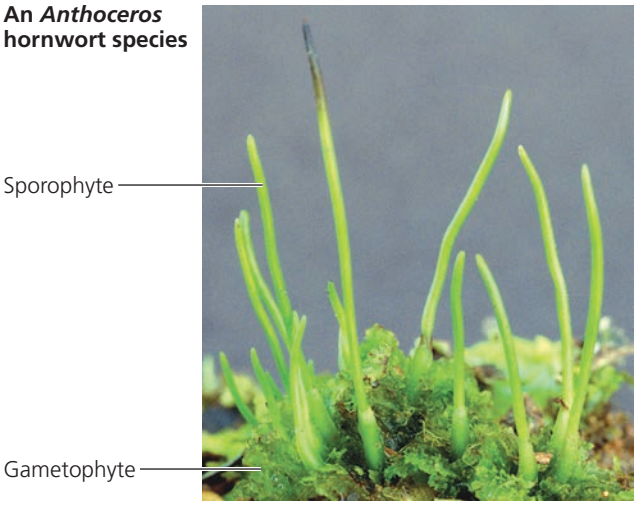
***Marchantia* sporophyte (LM)**

***Plagiochila deltoidea*, a "leafy" liverwort**

Hornworts (Phylum Anthocerophyta)

This phylum's common and scientific names (from the Greek keras, horn) refer to the long, tapered shape of the sporophyte. A typical sporophyte can grow to about 5 cm high. Unlike a liverwort or moss sporophyte, a hornwort sporophyte lacks a seta and consists only of a sporangium. The sporangium releases mature spores by splitting open, starting at the tip of the horn. The gametophytes, which are usually 1–2 cm in diameter, grow mostly horizontally and often have multiple sporophytes attached. Hornworts are frequently among the first species to colonize open areas with moist soils; a symbiotic relationship with nitrogen-fixing cyanobacteria contributes to their ability to do this (nitrogen is often in short supply in such areas).

An *Anthoceros* hornwort species



Mosses (Phylum Bryophyta)

Moss gametophytes, which range in height from less than 1 mm up to 60 cm, are less than 15 cm tall in most species. The familiar carpet of moss you observe consists mainly of gametophytes. The blades of their "leaves" are usually only one cell thick, but more complex "leaves" that have ridges coated with cuticle can be found on the common hairy-cap moss (Polytrichum) and its close relatives. Moss sporophytes are typically elongated and visible to the naked eye, with heights ranging up to about 20 cm. Though green and photosynthetic when young, they turn tan or brownish red when ready to release spores.

***Dicranum polysetum*, rugose fork-moss**

Many bryophyte species can increase the number of individuals in a local area through various methods of asexual reproduction. For example, some mosses reproduce asexually by forming *brood bodies*, small plantlets (as shown here) that detach from the parent plant and grow into new, genetically identical copies of their parent.

Bryophyte Sporophytes

The cells of bryophyte sporophytes contain plastids that are usually green and photosynthetic when the sporophytes are young. Even so, bryophyte sporophytes cannot live independently. A bryophyte sporophyte remains attached to its parental gametophyte throughout the sporophyte's lifetime, dependent on the gametophyte for supplies of sugars, amino acids, minerals, and water.

Bryophytes have the smallest sporophytes of all extant plant groups, consistent with the hypothesis that larger sporophytes evolved only later, in the vascular plants. A typical bryophyte sporophyte consists of a foot, a seta, and a sporangium. Embedded in the archegonium, the **foot** absorbs nutrients from the gametophyte. The **seta** (plural, *setae*), or stalk, conducts these materials to the sporangium, also called a **capsule**, which produces spores by meiosis.

Bryophyte sporophytes can produce enormous numbers of spores. A single moss capsule, for example, can generate over 5 million spores. In most mosses, the seta becomes elongated, enhancing spore dispersal by elevating the capsule. Typically, the upper part of the capsule features a ring of interlocking, tooth-like structures known as the **peristome** (see Figure 29.7). These "teeth" open under dry conditions and close again when it is moist. This allows moss spores to be discharged gradually, via periodic gusts of wind that can carry them long distances.

Moss and hornwort sporophytes are often larger and more complex than those of liverworts. For example, hornwort sporophytes, which superficially resemble grass blades, have a cuticle. Moss and hornwort sporophytes also have stomata, as do all vascular plants (but not liverworts).

Figure 29.8 shows some examples of gametophytes and sporophytes in the bryophyte phyla.

The Ecological and Economic Importance of Mosses

Wind dispersal of lightweight spores has distributed mosses throughout the world. These plants are particularly common and diverse in moist forests and wetlands. Some mosses colonize bare, sandy soil, where, researchers have found, they help retain nitrogen in the soil. In northern coniferous forests, species such as the feather moss *Pleurozium* harbor nitrogen-fixing cyanobacteria that increase the availability of nitrogen in the ecosystem. Other mosses inhabit such extreme environments as mountaintops, tundra, and deserts. Many mosses are able to live in very cold or dry habitats because they can survive the loss of most of their body water, then rehydrate when moisture is available. Few vascular plants can survive the same degree of desiccation. Moreover, phenolic compounds in moss cell walls absorb damaging levels of UV radiation present in deserts or at high altitudes.

One wetland moss genus, *Sphagnum*, or peat moss, is often a major component of deposits of partially decayed organic

Figure 29.9

Inquiry Can bryophytes help prevent landslides on tropical mountains?

Field Study Tropical lowland forests are poor in bryophytes. Higher up on the mountains, however, in so-called mossy forests or cloud forests (because they are often shrouded in clouds), the soil, tree trunks, and branches are all laden with bryophytes due to lower temperatures and higher moisture. The tree-dwelling bryophytes as well as other plants and lichens are called epiphytes. An example of a mossy forest on a tropical mountain is shown below. Tamás Pócs, a Hungarian botanist with much field experience in the tropics, studied the epiphytic biomass and its water-holding capacity in the mossy forests on the Uluguru Mountains in southeastern Tanzania, Africa.

Results In the mossy forests on the Uluguru Mountains, the total epiphyte mass was found to be approximately 13.7 tonnes dry weight per hectare. The epiphytic bryophytes alone were able to absorb nearly 30,000 L (30 tonnes) of water per hectare during a single rain after a relatively dry period.

Conclusion The thick bryophyte cover on the trees and soil acts as a giant sponge to significantly alleviate the effect of torrential tropical rains, considerably reducing soil erosion, occurrence of landslides, and leaching of nutrients on the mountain slopes. The bryophytes allow the water to filter gradually to the soil, and they regulate the level of agriculturally important watercourses.

Source: T. Pócs, Tropical forest bryophytes, in A. J. E. Smith (editor), *Bryophyte Ecology*, 59–104 (Chapman & Hall, 1982).

WHAT IF? ➤ *The mountain slopes below mossy forests are often in agricultural use—for example, as coffee plantations. What would happen to these plantations if the mossy forests were felled?*

material known as **peat** **(Figure 29.10a)**. Boggy regions with thick layers of peat are called peatlands. *Sphagnum* does not decay readily, in part because of phenolic compounds embedded in its cell walls. The low temperature, pH, and oxygen level of peatlands also inhibit decay of moss and other organisms in these boggy wetlands. As a result, some peatlands have preserved corpses for thousands of years **(Figure 29.10b)**.

Peat has long been a fuel source in Europe and Asia, and it is still harvested for fuel today, notably in Ireland and Canada. Peat moss is also useful as a soil conditioner and for packing plant roots during shipment because it has large dead cells that can absorb roughly 20 times the moss's weight in water.

Peatlands cover 3% of Earth's land surface and contain roughly 30% of the world's soil carbon: Globally, an estimated 450 billion tons of organic carbon is stored as peat. Current overharvesting of *Sphagnum*—primarily for use in peat-fired power stations—may contribute to global warming by releasing stored CO_2. In addition, if global temperatures continue to rise, the water levels of some peatlands are expected to drop. Such a change would expose peat to air and cause it to decompose, thereby releasing additional stored CO_2 and contributing further to global warming. The historical and expected future effects of *Sphagnum* on the global climate underscore the importance of preserving and managing peatlands.

Mosses may have a long history of affecting climate change. In the **Scientific Skills Exercise**, you will explore the question of whether they did so during the Ordovician period by contributing to the weathering of rocks.

Figure 29.10 ***Sphagnum*, or peat moss: a bryophyte with economic, ecological, and archaeological significance.**

(a) Peat being harvested from a peatland

(b) "Tollund Man," a bog mummy dating from 405–100 B.C.E. The acidic, oxygen-poor conditions produced by *Sphagnum* can preserve human or other animal bodies for thousands of years.

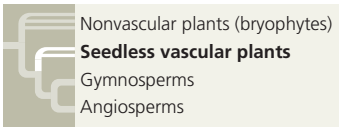
CONCEPT CHECK 29.2

1. How do bryophytes differ from other plants?
2. Explain why many bryophytes require a moist habitat to survive.
3. **MAKE CONNECTIONS ➤** Review the discussion of feedback regulation in Concept 1.1. Could effects of global warming on peatlands alter CO_2 concentrations in ways that result in negative or positive feedback? Explain.

For suggested answers, see Appendix A.

CONCEPT 29.3

Ferns and other seedless vascular plants were the first plants to grow tall



During the first 100 million years of plant evolution, bryophytes were prominent types of vegetation. But it is vascular plants that dominate most landscapes today. The earliest fossils of vascular plants date to 425 million years ago. These plants lacked seeds but had well-developed vascular systems, an evolutionary novelty that set the stage for vascular plants to grow taller than their bryophyte counterparts. As in bryophytes, however, the sperm of ferns and all other seedless vascular plants are flagellated and swim through a film of water to reach eggs. In part because of these swimming sperm, seedless vascular plants today are most common in damp environments.

Origins and Traits of Vascular Plants

Unlike the nonvascular plants, ancient relatives of vascular plants had branched sporophytes that were not dependent on gametophytes for nutrition **(Figure 29.11)**. Although these early plants were less than 20 cm tall, their branching enabled their bodies to become more complex and to have multiple sporangia. As plant bodies became more complex over time, competition for space and sunlight probably increased. As we'll see, that competition may have stimulated still more evolution in vascular plants, eventually leading to the formation of the first forests.

Early vascular plants had some derived traits of today's vascular plants, but they lacked roots and some other adaptations that evolved later. The main traits that characterize living vascular plants are life cycles with dominant sporophytes,

SCIENTIFIC SKILLS EXERCISE

Making Bar Graphs and Interpreting Data

Could Nonvascular Plants Have Caused Weathering of Rocks and Contributed to Climate Change During the Ordovician Period? The oldest traces of terrestrial plants are fossilized spores formed 470 million years ago. Between that time and the end of the Ordovician period 444 million years ago, the atmospheric CO_2 level dropped by half, and the climate cooled dramatically.

One possible cause of the drop in CO_2 during the Ordovician period is the breakdown, or weathering, of rock. As rock weathers, calcium silicate (Ca_2SiCO_3) is released and combines with CO_2 from the air, producing calcium carbonate ($CaCO_3$). Today, the roots of vascular plants increase rock weathering and mineral release by producing acids that break down rock and soil. Although nonvascular plants lack roots, they require the same mineral nutrients as vascular plants. Could nonvascular plants also increase the chemical weathering of rock? If so, they could have contributed to the decline in atmospheric CO_2 during the Ordovician. In this exercise, you will interpret data from a study of the effects of moss on releasing minerals from two types of rock.

How the Experiment Was Done The researchers set up experimental and control microcosms, or small artificial ecosystems, to measure mineral release from rocks. First, they placed rock fragments of volcanic origin, either granite or andesite, into small glass containers. Then they mixed water and macerated (chopped and crushed) moss of the species *Physcomitrella patens*. They added this mixture to the experimental microcosms (72 granite and 41 andesite). For the control microcosms (77 granite and 37 andesite), they filtered out the moss and just added the water. After 130 days, they measured the amounts of various minerals found in the water in the control microcosms and in the water and moss in the experimental microcosms.

Data from the Experiment The moss grew (increased its biomass) in the experimental microcosms. The table shows the mean amounts in micromoles (μmol) of several minerals measured in the water and the moss in the microcosms.

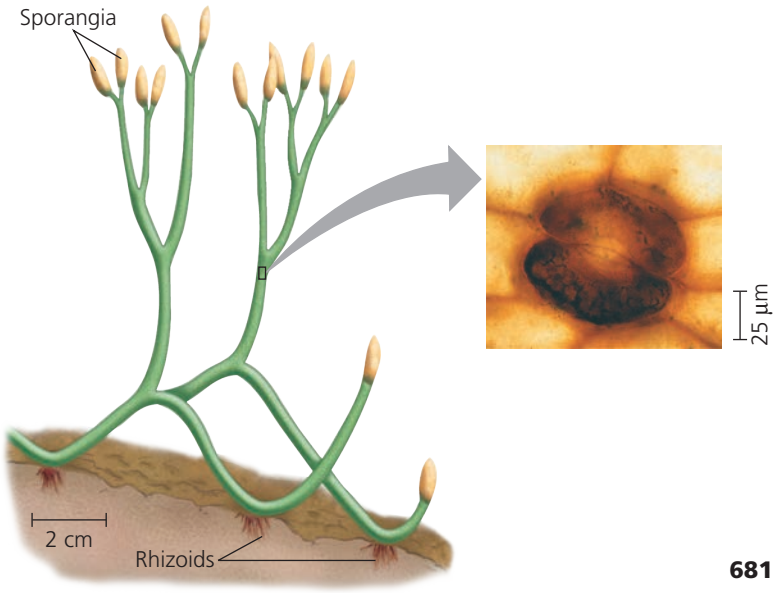
INTERPRET THE DATA

1. Why did the researchers add filtrate from which macerated moss had been removed to the control microcosms?
2. Make two bar graphs (for granite and andesite) comparing the mean amounts of each element weathered from rocks in the control and experimental microcosms. (Hint: For an experimental microcosm, what sum represents the total amount weathered from rocks?)
3. Overall, what is the effect of moss on chemical weathering of rock? Are the results similar or different for granite and andesite?
4. Based on their experimental results, the researchers added weathering of rock by nonvascular plants to simulation models of the Ordovician climate. The new models predicted decreased CO_2 levels and global cooling sufficient to produce the glaciations in the late Ordovician period. What assumptions did the researchers make in using results from their experiments in climate simulation models?
5. "Life has profoundly changed the Earth." Explain whether or not these experimental results support this statement.

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

	Ca^{2+} (μmol)		Mg^{2+} (μmol)		K^+ (μmol)	
	Granite	Andesite	Granite	Andesite	Granite	Andesite
Mean weathered amount released in water in the control microcosms	1.68	1.54	0.42	0.13	0.68	0.60
Mean weathered amount released in water in the experimental microcosms	1.27	1.84	0.34	0.13	0.65	0.64
Mean weathered amount taken up by moss in the experimental microcosms	1.09	3.62	0.31	0.56	1.07	0.28

Data from T. M. Lenton et al., First plants cooled the Ordovician, *Nature Geoscience* 5:86–89 (2012).



➤ **Figure 29.11 Sporophytes of *Aglaophyton major*, an ancient relative of living vascular plants.** This reconstruction from 405-million-year-old fossils exhibits dichotomous (Y-shaped) branching with sporangia at the ends of branches. Sporophyte branching characterizes living vascular plants but is lacking in living nonvascular plants (bryophytes). *Aglaophyton* had structures called rhizoids that anchored it to the ground. The inset shows a fossilized stoma of *A. major* (colorized LM).

transport in vascular tissues called xylem and phloem, and well-developed roots and leaves, including spore-bearing leaves called sporophylls.

Life Cycles with Dominant Sporophytes

As mentioned earlier, mosses and other bryophytes have life cycles dominated by gametophytes (see Figure 29.7). Fossil evidence suggests that a change began to develop in some of the earliest vascular plants, whose gametophytes and sporophytes were about equal in size. Further reductions in gametophyte size occurred among extant vascular plants; in these groups, the sporophyte generation is the larger and more complex form in the alternation of generations **(Figure 29.12)**. In ferns, for example, the familiar leafy plants are the sporophytes. You would have to get down on your hands and knees and search the ground carefully to find fern gametophytes, which are tiny structures that often grow on or just below the soil surface.

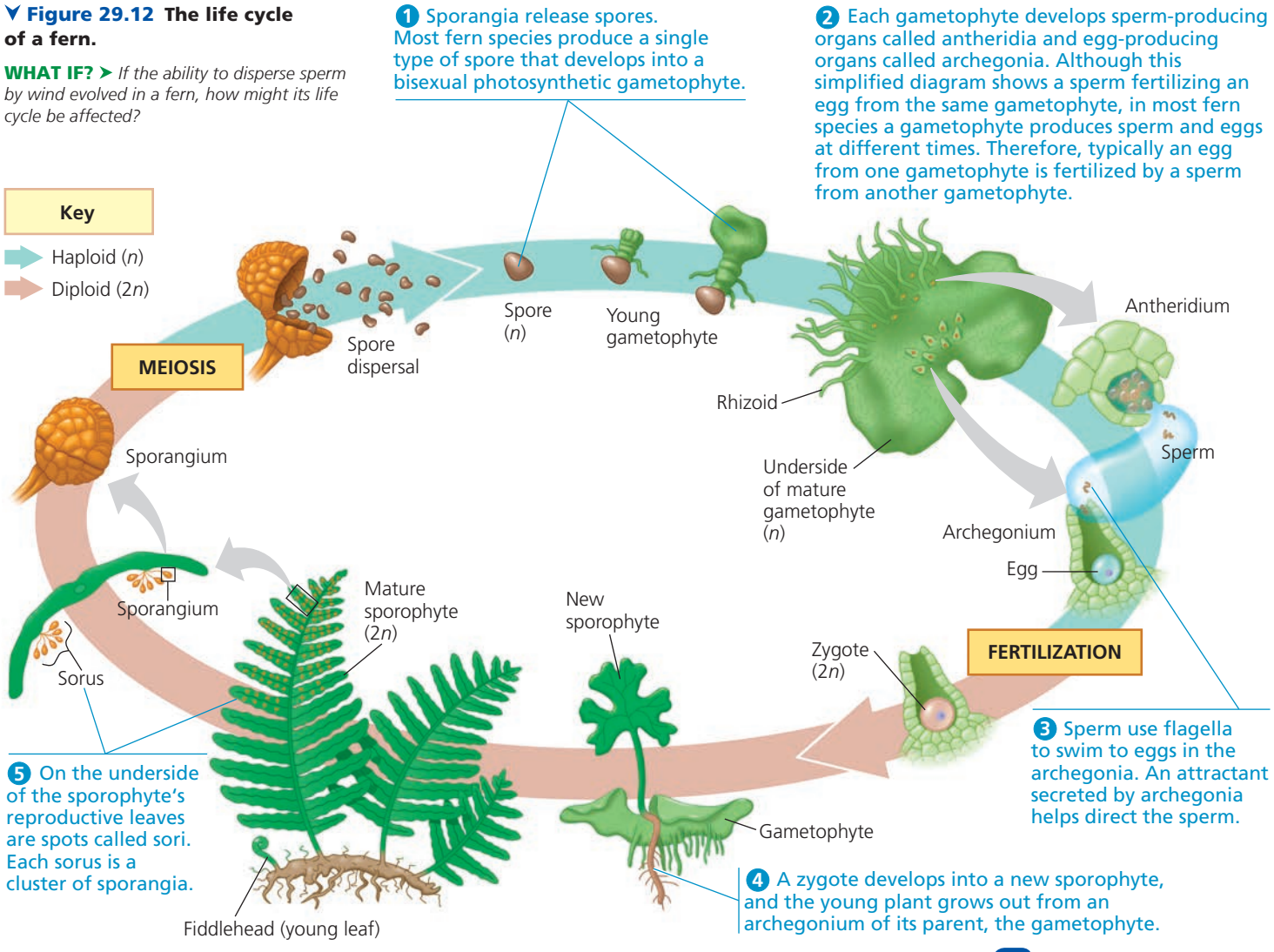
Transport in Xylem and Phloem

Vascular plants have two types of vascular tissue: xylem and phloem. **Xylem** conducts most of the water and minerals. The xylem of most vascular plants includes **tracheids**, tube-shaped cells that carry water and minerals up from the roots (see Figure 35.10). The water-conducting cells in vascular plants are *lignified*; that is, their cell walls are strengthened by the polymer **lignin**. The tissue called **phloem** has cells arranged into tubes that distribute sugars, amino acids, and other organic products (see Figure 35.10).

Lignified vascular tissue helped enable vascular plants to grow tall. Their stems became strong enough to provide support against gravity, and they could transport water and mineral nutrients high above the ground. Tall plants could also outcompete short plants for access to the sunlight needed for photosynthesis. In addition, the spores of tall plants could disperse farther than those of short plants, enabling tall species to colonize new environments rapidly. Overall, the ability to grow tall gave vascular plants a competitive edge over nonvascular plants, which typically are less than 5 cm in height. Competition among vascular plants also would have increased, leading to selection for taller growth forms—a process that eventually gave rise to the trees that formed the first forests about 385 million years ago.

Figure 29.12 The life cycle of a fern.

WHAT IF? ➤ *If the ability to disperse sperm by wind evolved in a fern, how might its life cycle be affected?*



MB Animation: Fern Life Cycle

Evolution of Roots

Vascular tissue also provides benefits below ground. Instead of the rhizoids seen in bryophytes, roots evolved in the sporophytes of almost all vascular plants. **Roots** are organs that absorb water and nutrients from the soil. Roots also anchor vascular plants to the ground, hence allowing the shoot system to grow taller.

Root tissues of living plants closely resemble stem tissues of early vascular plants preserved in fossils. This suggests that roots may have evolved from the lowest belowground portions of stems in ancient vascular plants. It is unclear whether roots evolved only once in the common ancestor of all vascular plants or independently in different lineages. Although the roots of living members of these lineages of vascular plants share many similarities, fossil evidence hints at convergent evolution. The oldest fossils of lycophytes, for example, already displayed simple roots 400 million years ago, when the ancestors of ferns and seed plants still had none. Studying genes that control root development in different vascular plant species may help resolve this question.

Evolution of Leaves

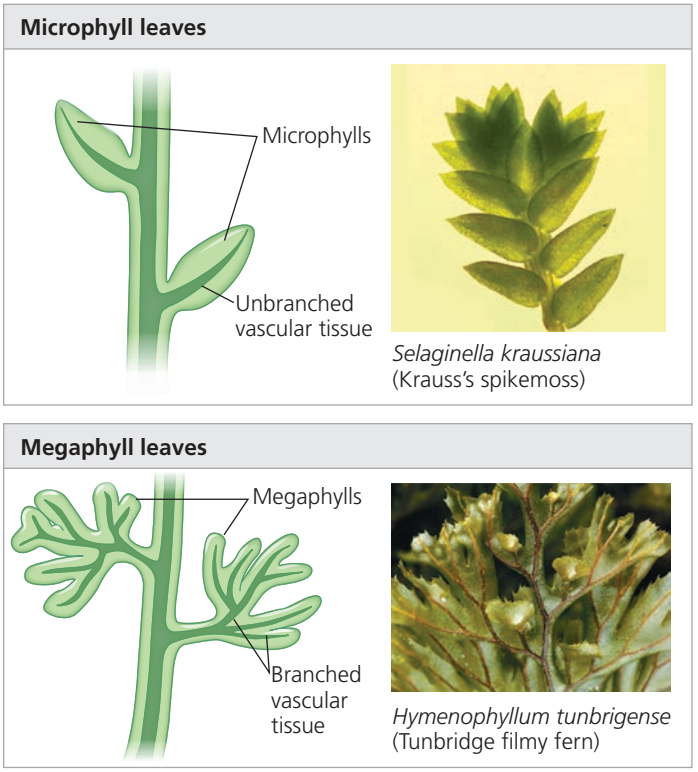
Leaves are structures that serve as the primary photosynthetic organ of vascular plants. In terms of size and complexity, leaves can be classified as either microphylls or megaphylls **(Figure 29.13)**. All of the lycophytes—and only the lycophytes—have **microphylls**, small, often spine-shaped leaves supported by a single strand of vascular tissue. Almost all other vascular plants have **megaphylls**, leaves with a highly branched vascular system; a few species have reduced leaves that appear to have evolved from megaphylls. Megaphylls are typically larger than microphylls and therefore support greater photosynthetic productivity than microphylls. Microphylls first appear in the fossil record 410 million years ago, but megaphylls do not emerge until about 370 million years ago, toward the end of the Devonian period.

Sporophylls and Spore Variations

One milestone in the evolution of plants was the emergence of **sporophylls**, modified leaves that bear sporangia. Sporophylls vary greatly in structure. For example, fern sporophylls produce clusters of sporangia known as **sori** (singular, *sorus*), usually on the undersides of the sporophylls (see Figure 29.12). In many lycophytes and in most gymnosperms, groups of sporophylls form cone-like structures called **strobili**. The sporophylls of angiosperms are called *carpels* and *stamens* (see Figure 30.8).

Most seedless vascular plant species are **homosporous**: They have one type of sporophyll bearing one type of sporangium that produces one type of spore, which typically develops into a bisexual gametophyte, as in most ferns. In contrast, a **heterosporous** species has two types of sporophylls: megasporophylls and microsporophylls. Megasporophylls have megasporangia, which produce **megaspores**, spores that develop into female gametophytes. Microsporophylls have microsporangia, which produce **microspores**, smaller spores that develop into male gametophytes. All seed plants and a few seedless vascular plants are heterosporous. The following diagram compares the two conditions:

Figure 29.13 Microphyll and megaphyll leaves.



Homosporous spore production
(most seedless vascular plants)

Sporangium on sporophyll → Single type of spore → Typically a bisexual gametophyte → Eggs / Sperm

Heterosporous spore production
(all seed plants)

Megasporangium on megasporophyll → Megaspore → Female gametophyte → Eggs

Microsporangium on microsporophyll → Microspore → Male gametophyte → Sperm

Classification of Seedless Vascular Plants

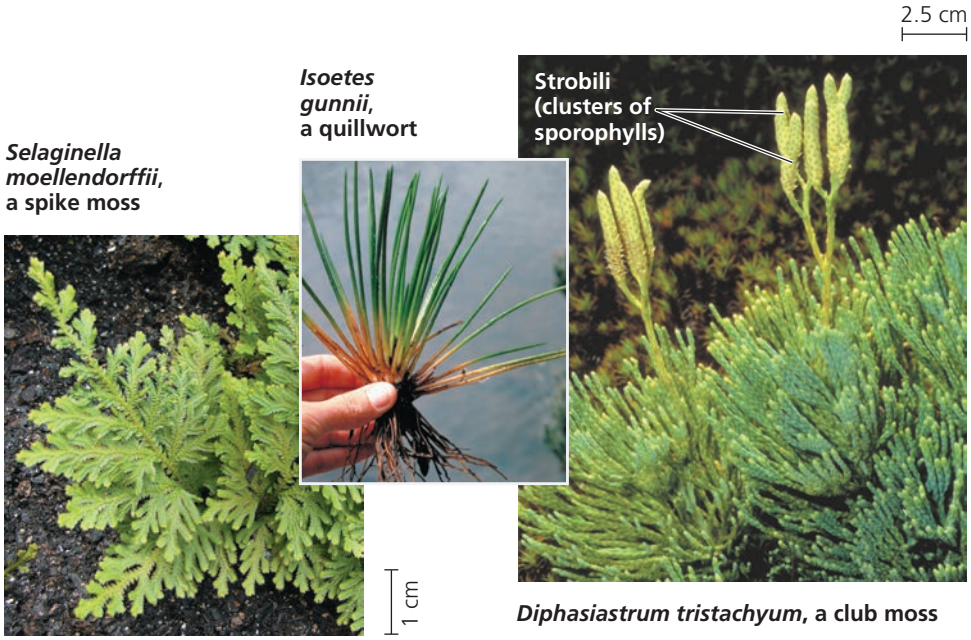
As we noted earlier, biologists recognize two clades of living seedless vascular plants: the lycophytes (phylum Lycophyta) and the monilophytes (phylum Monilophyta). The lycophytes include the club mosses, the spikemosses, and the quillworts. The monilophytes include the ferns, the horsetails, and the whisk ferns and their relatives. Although ferns, horsetails,

Figure 29.14 Exploring Seedless Vascular Plant Diversity

Lycophytes (Phylum Lycophyta)

Many lycophytes grow on tropical trees as epiphytes, plants that use other plants as a substrate but are not parasites. Other species grow on temperate forest floors. In some species, the tiny gametophytes live above ground and are photosynthetic. Others live below ground, nurtured by symbiotic fungi.

Sporophytes have upright stems with many small leaves, as well as ground-hugging stems that produce dichotomously branching roots. Spike mosses are usually relatively small and often grow horizontally. In many club mosses and spike mosses, sporophylls are clustered into club-shaped cones (strobili). Quillworts, named for their leaf shape, form a single genus whose members live in marshy areas or as submerged aquatic plants. Club mosses are all homosporous, whereas spike mosses and quillworts are all heterosporous. The spores of club mosses are released in clouds and are so rich in oil that magicians and photographers once ignited them to create smoke or flashes of light.



Diphasiastrum tristachyum, a club moss

Monilophytes (Phylum Monilophyta)

Matteuccia struthiopteris (ostrich fern)

Ferns

Unlike the lycophytes, ferns have megaphylls (see Figure 29.13). The sporophytes typically have horizontal stems that give rise to large leaves called fronds, often divided into leaflets. A frond grows as its coiled tip, the fiddlehead, unfurls.

Almost all species are homosporous. The gametophyte in some species shrivels and dies after the young sporophyte detaches itself. In most species, sporophytes have stalked sporangia with springlike devices that catapult spores several meters. Airborne spores can be carried far from their origin. Some species produce more than a trillion spores in a plant's lifetime.

Equisetum telmateia, giant horsetail

Horsetails

The group's name refers to the brushy appearance of the stems, which have a gritty texture that made them historically useful as "scouring rushes" for pots and pans. Some species have separate fertile (cone-bearing) and vegetative stems. Horsetails are homosporous, with cones releasing spores that typically give rise to bisexual gametophytes.

Horsetails are also called arthrophytes ("jointed plants") because their stems have joints. Rings of small leaves or branches emerge from each joint, but the stem is the main photosynthetic organ. Large air canals carry oxygen to the roots, which often grow in waterlogged soil.

Psilotum nudum, a whisk fern

Whisk Ferns and Relatives

Like primitive vascular plant fossils, the sporophytes of whisk ferns (genus *Psilotum*) have dichotomously branching stems but no roots. Stems have scalelike outgrowths that lack vascular tissue and may have resulted from the evolutionary reduction of leaves. Each yellow knob on a stem consists of three fused sporangia. Species of the genus *Tmesipteris* closely related to whisk ferns and found only in the South Pacific, also lack roots but have small, leaflike outgrowths in their stems, giving them a vine-like appearance. Both genera are homosporous, with spores giving rise to bisexual gametophytes that grow underground and are only about a centimeter long.

and whisk ferns differ greatly in appearance, recent anatomical and molecular comparisons provide convincing evidence that these three groups make up a clade. Accordingly, many systematists now classify them together as the phylum Monilophyta, as we do in this chapter. Others refer to these groups as three separate phyla within a clade. **Figure 29.14** describes the two main groups of seedless vascular plants.

Phylum Lycophyta: Club Mosses, Spikemosses, and Quillworts

Present-day species of lycophytes are relicts of a far more impressive past. By the Carboniferous period (359–299 million years ago), the lycophyte evolutionary lineage included small herbaceous plants and giant trees with diameters of more than 2 m and heights of more than 40 m. The giant lycophyte trees thrived for millions of years in moist swamps, but their diversity declined when Earth's climate became drier during the Permian period (299–252 million years ago). The small lycophytes survived, represented today by about 1,200 species. Though some are commonly called club mosses and spikemosses, they are not true mosses (which, as discussed earlier, are nonvascular plants).

Phylum Monilophyta: Ferns, Horsetails, and Whisk Ferns and Relatives

Ferns radiated extensively from their Devonian origins and grew alongside lycophyte trees and horsetails in the great Carboniferous swamp forests. Today, ferns are by far the most widespread seedless vascular plants, numbering more than 12,000 species. Though most diverse in the tropics, many ferns thrive in temperate forests, and some species are even adapted to arid habitats.

As mentioned earlier, ferns and other monilophytes are more closely related to seed plants than to lycophytes. As a result, monilophytes and seed plants share traits that are not found in lycophytes, including megaphyll leaves and roots that can branch at various points along the length of an existing root. In lycophytes, by contrast, roots branch only at the growing tip of the root, forming a Y-shaped structure.

The monilophytes called horsetails were very diverse during the Carboniferous period, some growing as tall as 15 m. Today, only 15 species survive as a single, widely distributed genus, *Equisetum*, often found in marshy places and along streams.

Psilotum (whisk ferns) and a closely related genus, *Tmesipteris*, form a clade consisting mainly of tropical epiphytes. Plants in these two genera, the only vascular plants lacking true roots, once were called "living fossils" because of their resemblance to fossils of ancient relatives of living vascular plants (see Figures 29.11 and 29.14). However, much evidence, including analyses of DNA sequences and sperm structure, indicates that the genera *Psilotum* and *Tmesipteris* are closely related to ferns. This hypothesis suggests that their ancestor's true roots were lost during evolution. Today, plants in these two genera absorb water and nutrients through numerous absorptive rhizoids.

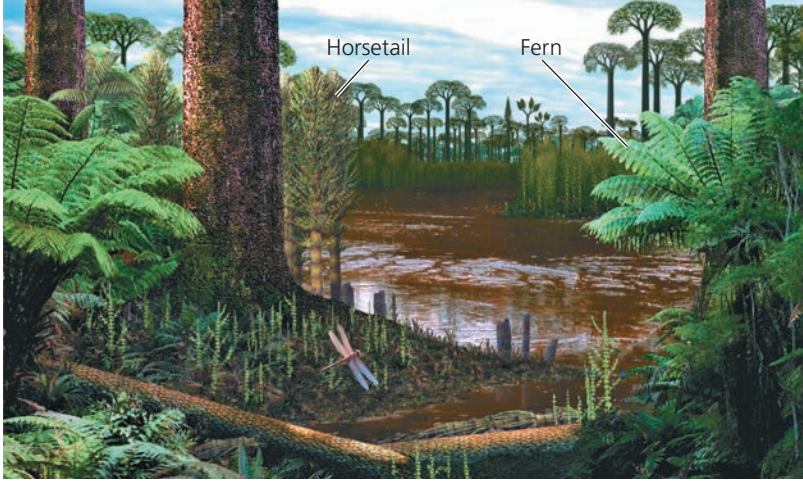


Figure 29.15 Artist's conception of a Carboniferous forest based on fossil evidence. Lycophyte trees, with trunks covered with small leaves, thrived in the "coal forests" of the Carboniferous, along with giant ferns and horsetails.

The Significance of Seedless Vascular Plants

The ancestors of living lycophytes, horsetails, and ferns, along with their extinct seedless vascular relatives, grew to great heights during the Devonian and early Carboniferous, forming the first forests **(Figure 29.15)**. How did their dramatic growth affect Earth and its other life?

One major effect was that early forests contributed to a large drop in CO_2 levels during the Carboniferous period, causing global cooling that resulted in widespread glacier formation. The trees of early forests contributed to this drop in CO_2 levels in part by the actions of their roots. The roots of vascular plants secrete acids that break down rocks, thereby increasing the rate at which calcium and magnesium are released from rocks into the soil. These chemicals react with carbon dioxide dissolved in rain water, forming compounds that ultimately wash into the oceans, where they are incorporated into rocks (calcium or magnesium carbonates). The net effect of these processes—which were accelerated by plants—is that CO_2 removed from the air is stored in marine rocks. Although carbon stored in these rocks can be returned to the atmosphere, it typically takes millions of years for this to occur (as when geological uplift brings the rocks to the surface, exposing them to erosion).

In addition, the seedless vascular plants that formed the first forests eventually became coal, again removing CO_2 from the atmosphere for long periods of time. In the stagnant waters of Carboniferous swamps, the dead bodies of early trees did not completely decay. This organic material turned to thick layers of peat, later covered by the sea. Marine sediments piled on top, and over millions of years, heat and pressure converted the peat to coal. In fact, Carboniferous coal deposits are the most extensive ever formed. Coal was crucial to the Industrial

Revolution, and people worldwide still burn 6 billion tons a year. It is ironic that coal, formed from plants that contributed to a global cooling, now contributes to global warming by returning carbon to the atmosphere (see Figure 56.29).

Growing along with the seedless plants in Carboniferous swamps were primitive seed plants. Though seed plants were not dominant at that time, they rose to prominence after the swamps began to dry up at the end of the Carboniferous period. The next chapter traces the origin and diversification of seed plants, continuing our story of adaptation to life on land.

CONCEPT CHECK 29.3

1. List the key derived traits found in monilophytes and seed plants, but not in lycophytes.
2. How do the main similarities and differences between seedless vascular plants and nonvascular plants affect function in these plants?
3. **MAKE CONNECTIONS ➤** In Figure 29.12, if fertilization occurred between gametes from one gametophyte, how would this affect the production of genetic variation from sexual reproduction? See Concept 13.4.

For suggested answers, see Appendix A.

29 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

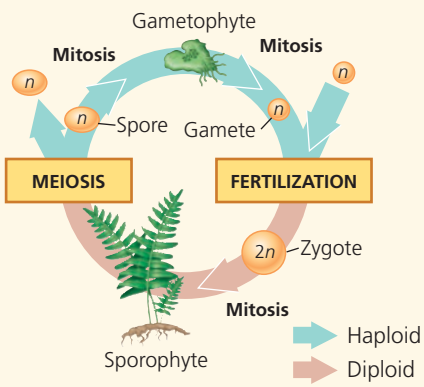
SUMMARY OF KEY CONCEPTS

CONCEPT 29.1

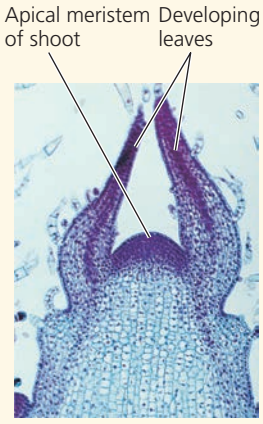
VOCAB SELF-QUIZ goo.gl/Rn5Uax

Plants evolved from green algae *(pp. 671–676)*

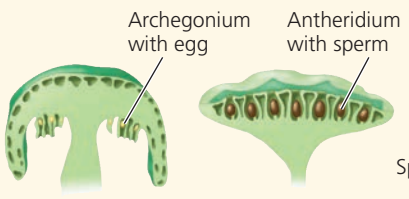
- Morphological and biochemical traits, as well as similarities in nuclear and chloroplast genes, indicate that certain groups of charophytes are the closest living relatives of plants.
- A protective layer of **sporopollenin** and other traits allow charophytes to tolerate occasional drying along the edges of ponds and lakes. Such traits may have enabled the algal ancestors of plants to survive in terrestrial conditions, opening the way to the colonization of dry land.
- Derived traits that distinguish plants from charophytes, their closest algal relatives, include **cuticles, stomata,** multicellular dependent embryos, and the four shown here:



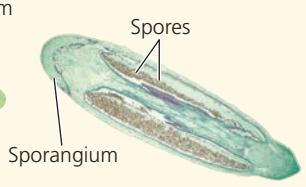
1 Alternation of generations



2 Apical meristems



3 Multicellular gametangia



4 Walled spores in sporangia

- Fossils show that plants arose more than 470 million years ago. Subsequently, plants diverged into several major groups, including nonvascular plants (**bryophytes**); **seedless vascular plants**, such as **lycophytes** and ferns; and the two groups of seed plants: **gymnosperms** and **angiosperms**.

? *Draw a phylogenetic tree illustrating our current understanding of plant phylogeny; label the common ancestor of plants and the origins of multicellular gametangia, vascular tissue, and seeds.*

CONCEPT 29.2

Mosses and other nonvascular plants have life cycles dominated by gametophytes *(pp. 676–680)*

- Lineages leading to the three extant clades of nonvascular plants, or **bryophytes—liverworts, mosses**, and **hornworts**—diverged from other plants early in plant evolution.
- In bryophytes, the dominant generation consists of haploid **gametophytes**, such as those that make up a carpet of moss. **Rhizoids** anchor gametophytes to the substrate on which they grow. The flagellated sperm produced by **antheridia** require a film of water to travel to the eggs in the **archegonia**.
- The diploid stage of the life cycle—the **sporophytes**—grow out of archegonia and are attached to the gametophytes and dependent on them for nourishment. Smaller and simpler than vascular plant sporophytes, they typically consist of a **foot, seta** (stalk), and **sporangium**.
- *Sphagnum*, or peat moss, is common in large regions known as peatlands and has many practical uses, including as a fuel.

? *Summarize the ecological importance of mosses.*

CONCEPT 29.3

Ferns and other seedless vascular plants were the first plants to grow tall *(pp. 680–686)*

- Fossils of the forerunners of today's vascular plants date back about 425 million years and show that these small plants had independent, branching sporophytes and a vascular system.
- Over time, other derived traits of living vascular plants arose, such as a life cycle with dominant sporophytes, lignified vascular tissue, well-developed **roots** and **leaves**, and **sporophylls**.
- Seedless vascular plants include the **lycophytes** (phylum Lycophyta: club mosses, spikemosses, and quillworts) and

the **monilophytes** (phylum Monilophyta: ferns, horsetails, and whisk ferns and relatives). Current evidence indicates that seedless vascular plants, like bryophytes, do not form a clade.
- Ancient lineages of lycophytes included both small herbaceous plants and large trees. Present-day lycophytes are small herbaceous plants.
- Seedless vascular plants formed the earliest forests about 385 million years ago. Their growth may have contributed to a major global cooling that took place during the Carboniferous period. The decaying remnants of the first forests eventually became coal.

What trait(s) allowed vascular plants to grow tall, and why might increased height have been advantageous?

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–5 can be found in the Study Area in MasteringBiology.

6. Identify each of the following structures as haploid or diploid.
(A) sporophyte
(B) spore
(C) gametophyte
(D) zygote

PRACTICE TEST
goo.gl/iAsVgL

7. EVOLUTION CONNECTION

DRAW IT Draw a phylogenetic tree that represents our current understanding of evolutionary relationships between a moss, a gymnosperm, a lycophyte, and a fern. Use a charophyte alga as the outgroup. (See Figure 22.5 to review phylogenetic trees.) Label each branch point of the phylogeny with at least one derived character unique to the clade descended from the common ancestor represented by the branch point.

8. SCIENTIFIC INQUIRY

INTERPRET THE DATA The feather moss *Pleurozium schreberi* harbors species of symbiotic nitrogen-fixing bacteria. Scientists studying this moss in northern forests found that the percentage of the ground surface "covered" by the moss increased from about 5% in forests that burned 35 to 41 years ago to about 70% in forests that burned 170 or more years ago. From mosses growing in these forests, they also obtained the following data on nitrogen fixation:

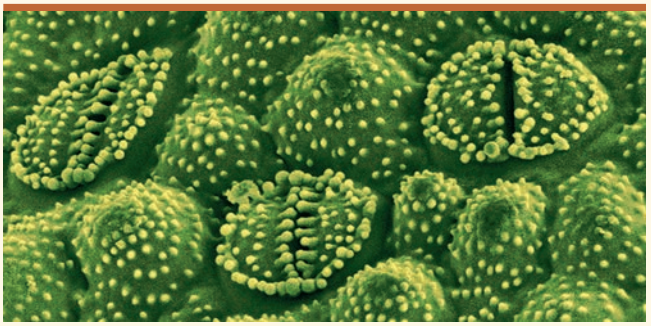
Age (years after fire)	N fixation rate [kg N/(ha · yr)]
35	0.001
41	0.005
78	0.08
101	0.3
124	0.9
170	2.0
220	1.3
244	2.1
270	1.6
300	3.0
355	2.3

Data from O. Zackrisson et al., Nitrogen fixation increases with successional age in boreal forests, *Ecology* 85:3327–3334 (2006).

(a) Use the data to draw a line graph, with age on the *x*-axis and the nitrogen fixation rate on the *y*-axis.
(b) Along with the nitrogen added by nitrogen fixation, about 1 kg of nitrogen per hectare per year is deposited into northern forests from the atmosphere as rain and small particles. Evaluate the extent to which *Pleurozium* affects nitrogen availability in northern forests of different ages.

9. WRITE ABOUT A THEME: INTERACTIONS Giant lycophyte trees had microphylls, whereas ferns and seed plants have megaphylls. Write a short essay (100–150 words) describing how a forest of lycophyte trees may have differed from a forest of large ferns or seed plants. In your answer, consider how the type of forest may have affected interactions among small plants growing beneath the tall ones.

10. SYNTHESIZE YOUR KNOWLEDGE



These stomata are from the leaf of a common horsetail. Describe how stomata and other adaptations facilitated life on land and ultimately led to the formation of the first forests.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

30 Seed Plants

Figure 30.1 How could these plants have reached this remote location?

KEY CONCEPTS

30.1 Seeds and pollen grains are key adaptations for life on land

30.2 Gymnosperms bear "naked" seeds, typically on cones

30.3 The reproductive adaptations of angiosperms include flowers and fruits

30.4 Human welfare depends on seed plants

Fireweed seed

Transforming the World

On May 18, 1980, Mount St. Helens erupted with a force 500 times that of the Hiroshima atomic bomb. Traveling at over 300 miles per hour, the blast destroyed hundreds of hectares of forest, leaving the region covered in ash and devoid of visible life. Within a few years, however, plants such as fireweed (*Chamerion angustifolium*) had colonized the barren landscape **(Figure 30.1)**.

Fireweed and other early arrivals reached the blast zone as seeds. A **seed** consists of an embryo and its food supply, surrounded by a protective coat. When mature, seeds are dispersed from their parent by wind or other means, enabling them to colonize distant locations.

Plants not only have affected the recovery of regions such as Mount St. Helens but also have transformed Earth. Continuing the saga of how this occurred, this chapter follows the emergence and diversification of the group to which fireweed belongs, the seed plants. Fossils and comparative studies of living plants offer clues about the origin of seed plants some 360 million years ago. As this new group became established, it dramatically altered the course of plant evolution. Indeed, seed plants have become the dominant producers on land, and they make up the vast majority of plant biodiversity today.

In this chapter, we will first examine the general features of seed plants. Then we will look at their evolutionary history and enormous impact on human society.

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

CONCEPT 30.1

Seeds and pollen grains are key adaptations for life on land

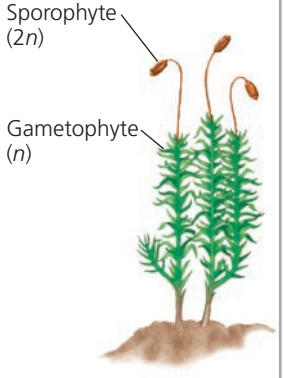
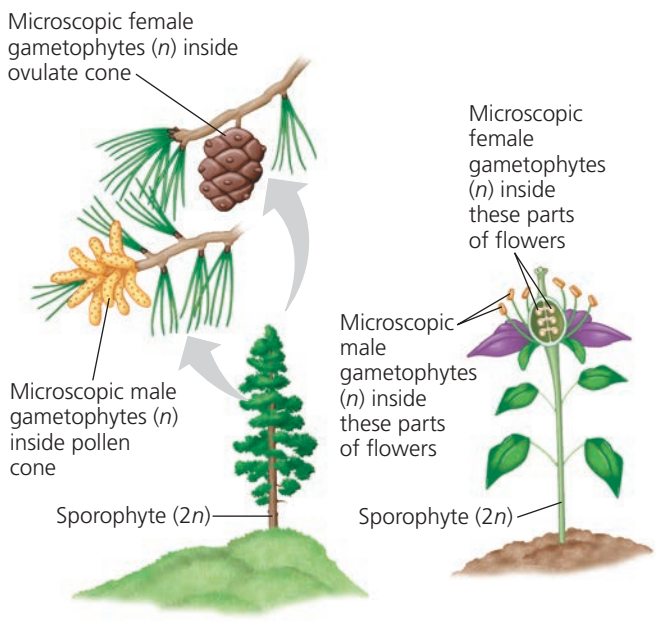
We begin with an overview of terrestrial adaptations that seed plants added to those already present in nonvascular plants (bryophytes) and seedless vascular plants (see Concept 29.1). In addition to seeds, all seed plants have reduced gametophytes, heterospory, ovules, and pollen. As we'll see, these adaptations helped seed plants cope with conditions such as drought and exposure to ultraviolet (UV) radiation in sunlight. They also freed seed plants from requiring water for fertilization, enabling reproduction under a broader range of conditions than in seedless plants.

Advantages of Reduced Gametophytes

Mosses and other bryophytes have life cycles dominated by gametophytes, whereas ferns and other seedless vascular plants have sporophyte-dominated life cycles. The evolutionary trend of gametophyte reduction continued further in the vascular plant lineage that led to seed plants. While the gametophytes of seedless vascular plants are visible to the naked eye, the gametophytes of most seed plants are microscopic.

This miniaturization allowed for an important evolutionary innovation in seed plants: Their tiny gametophytes can develop from spores retained within the sporangia of the parental sporophyte. This arrangement can protect the gametophytes from environmental stresses. For example, the moist reproductive tissues of the sporophyte shield the gametophytes from UV radiation and protect them from drying out. This relationship also enables the developing gametophytes to obtain nutrients from the parental sporophyte. In contrast, the free-living gametophytes of seedless vascular plants must fend for themselves. **Figure 30.2** provides an overview of the gametophyte-sporophyte relationships in nonvascular plants, seedless vascular plants, and seed plants.

Figure 30.2 Gametophyte-sporophyte relationships in different plant groups.

	PLANT GROUP		
	Mosses and other nonvascular plants	**Ferns and other seedless vascular plants**	**Seed plants (gymnosperms and angiosperms)**
Gametophyte	Dominant	Reduced, independent (photosynthetic and free-living)	Reduced (usually microscopic), dependent on surrounding sporophyte tissue for nutrition
Sporophyte	Reduced, dependent on gametophyte for nutrition	Dominant	Dominant
Example			**Gymnosperm** **Angiosperm** 

MAKE CONNECTIONS ➤ *In seed plants, how might retaining the gametophyte within the sporophyte affect embryo fitness? (See Concepts 17.5, 23.1, and 23.4 to review mutagens, mutations, and fitness.)*

Heterospory: The Rule Among Seed Plants

You read in Concept 29.3 that most seedless plants are *homosporous*—they produce one kind of spore, which usually gives rise to a bisexual gametophyte. Ferns and other close relatives of seed plants are homosporous, suggesting that seed plants had homosporous ancestors. At some point, seed plants or their ancestors became *heterosporous*, producing two kinds of spores: Megasporangia on modified leaves called megasporophylls produce *megaspores* that give rise to female gametophytes, and microsporangia on modified leaves called microsporophylls produce *microspores* that give rise to male gametophytes. Each megasporangium has one megaspore, whereas each microsporangium has many microspores.

As noted previously, the miniaturization of seed plant gametophytes probably contributed to the great success of this clade. Next, we'll look at the development of the female gametophyte within an ovule and the development of the male gametophyte in a pollen grain. Then we'll follow the transformation of a fertilized ovule into a seed.

Ovules and Production of Eggs

Although a few species of seedless plants are heterosporous, seed plants are unique in retaining the megasporangium within the parent sporophyte. A layer of sporophyte tissue called **integument** envelops and protects the megasporangium. Gymnosperm megasporangia are surrounded by one integument, whereas those in angiosperms usually have two integuments. The whole structure—megasporangium, megaspore, and their integument(s)—is called an **ovule** **(Figure 30.3a)**. Inside each ovule (from the Latin *ovulum*, little egg), a female gametophyte develops from a megaspore and produces one or more eggs.

Pollen and Production of Sperm

A microspore develops into a **pollen grain** that consists of a male gametophyte enclosed within the pollen wall. (The wall's outer layer is made of molecules secreted by sporophyte cells, so we refer to the male gametophyte as being *in* the pollen grain, not *equivalent to* the pollen grain.) Sporopollenin in the pollen wall protects the pollen grain as it is transported by wind or by hitchhiking on an animal. The transfer of pollen to the part of a seed plant that contains the ovules is called **pollination**. If a pollen grain germinates (begins growing), it gives rise to a pollen tube that discharges sperm into the female gametophyte within the ovule, as shown in **Figure 30.3b**.

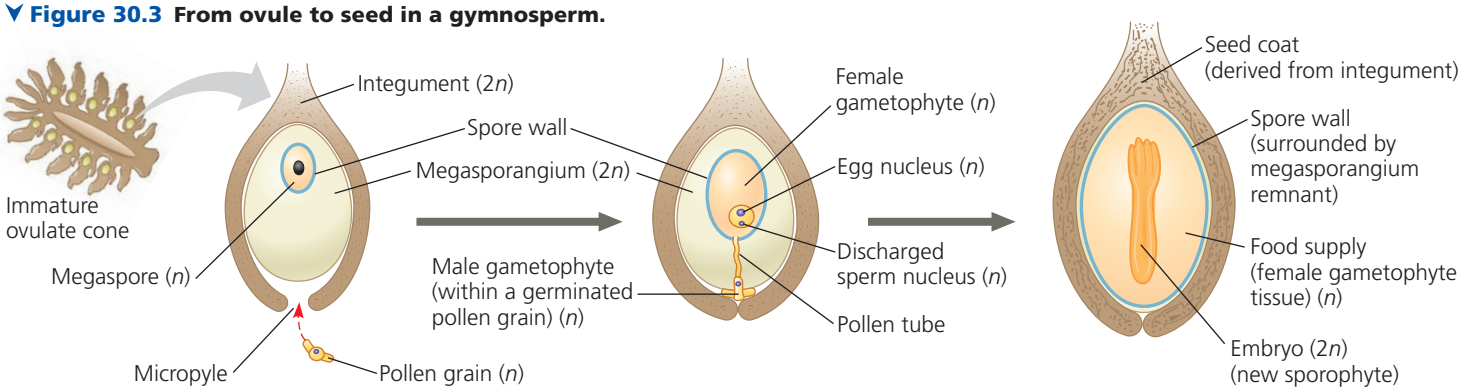
In nonvascular plants and seedless vascular plants such as ferns, free-living gametophytes release flagellated sperm that swim through a film of water to reach eggs. Given this requirement, it is not surprising that many of these species live in moist habitats. But a pollen grain can be carried by wind or animals, eliminating the dependence on water for sperm transport. The ability of seed plants to transfer sperm without water likely contributed to their colonization of dry habitats. The sperm of seed plants also do not require motility because they are carried to the eggs by pollen tubes. The sperm of some gymnosperm species retain the ancient flagellated condition, but flagella have been lost in the sperm of most gymnosperms and all angiosperms.

The Evolutionary Advantage of Seeds

If a sperm fertilizes an egg of a seed plant, the zygote grows into a sporophyte embryo. As shown in **Figure 30.3c**, the ovule develops into a seed: the embryo, with a food supply, packaged in a protective coat derived from the integument(s).

Until the advent of seeds, the spore was the only protective stage in any plant life cycle. Moss spores, for example, may

Figure 30.3 From ovule to seed in a gymnosperm.



(a) Unfertilized ovule. In this longitudinal section through the ovule of a pine (a gymnosperm), a fleshy megasporangium is surrounded by a protective layer of tissue called an integument. The micropyle, the only opening through the integument, allows entry of a pollen grain.

(b) Fertilized ovule. A megaspore develops into a female gametophyte, which produces an egg. The pollen grain, which had entered through the micropyle, contains a male gametophyte. The male gametophyte develops a pollen tube that discharges sperm, thereby fertilizing the egg.

(c) Gymnosperm seed. Fertilization initiates the transformation of the ovule into a seed, which consists of a sporophyte embryo, a food supply, and a protective seed coat derived from the integument. The megasporangium dries out and collapses.

VISUAL SKILLS ➤ *Based on this diagram, a gymnosperm seed contains cells from how many different plant generations? Identify the cells and whether each is haploid or diploid.*

MB **Figure Walkthrough**

SCIENTIFIC SKILLS EXERCISE

Using Natural Logarithms to Interpret Data

How Long Can Seeds Remain Viable in Dormancy?
Environmental conditions can vary greatly over time, and they may not be favorable for germination when seeds are produced. One way that plants cope with such variation is through seed dormancy. Under favorable conditions, seeds of some species can germinate after many years of dormancy.

One unusual opportunity to test how long seeds can remain viable occurred when seeds from date palm trees (*Phoenix dactylifera*) were discovered under the rubble of a 2,000-year-old fortress near the Dead Sea. As you saw in the Chapter 2 Scientific Skills Exercise and Concept 25.2, scientists use radiometric dating to estimate the ages of fossils and other old objects. In this exercise, you will estimate the ages of three of these ancient seeds by using natural logarithms.

How the Experiment Was Done Scientists measured the fraction of carbon-14 that remained in three ancient date palm seeds: two that were not planted and one that was planted and germinated. For the germinated seed, the scientists used a seed coat fragment found clinging to a root of the seedling. (The seedling grew into the plant in the photo.)

Data from the Experiment This table shows the fraction of carbon-14 remaining from the three ancient date palm seeds.

	Fraction of Carbon-14 Remaining
Seed 1 (not planted)	0.7656
Seed 2 (not planted)	0.7752
Seed 3 (germinated)	0.7977

INTERPRET THE DATA

A logarithm is the power to which a base is raised to produce a given number x. For example, if the base is 10 and $x = 100$, the logarithm of 100 equals 2 (because $10^2 = 100$). A natural logarithm (ln) is the logarithm of a number x to the base e, where e is about 2.718. Natural logarithms are useful in calculating rates of some natural processes, such as radioactive decay.

1. The equation $F = e^{-kt}$ describes the fraction F of an original isotope remaining after a period of t years; the exponent is negative because it refers to a *decrease* over time. The constant k provides a measure of how rapidly the original isotope decays. For the decay of carbon-14 to nitrogen-14, $k = 0.00012097$. To find t, rearrange the equation by following these steps: (a) Take the natural logarithm of both sides of the equation: $\ln(F) = \ln(e^{-kt})$. Rewrite the right side of this equation by applying the following rule: $\ln(e^x) = x \ln(e)$. (b) Since $\ln(e) = 1$, simplify the equation. (c) Now solve for t and write the equation in the form "$t =$ _____."
2. Using the equation you developed, the data from the table, and a calculator, estimate the ages of seed 1, seed 2, and seed 3.
3. Why do you think there was more carbon-14 in the germinated seed?

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Data from S. Sallon et al., Germination, genetics, and growth of an ancient date seed, *Science* 320:1464 (2008).

survive even if the local environment becomes too cold, too hot, or too dry for the mosses themselves to live. Their tiny size enables the spores to be dispersed in a dormant state to a new area, where they can germinate and give rise to new moss gametophytes if and when conditions are favorable enough for them to break dormancy. Spores were the main way that mosses, ferns, and other seedless plants spread over Earth for the first 100 million years of plant life on land.

Although mosses and other seedless plants continue to be very successful today, seeds represent a major evolutionary innovation that contributed to the opening of new ways of life for seed plants. What advantages do seeds provide over spores? Spores are usually single-celled, whereas seeds are multicellular, consisting of an embryo protected by a layer of tissue, the seed coat. A seed can remain dormant for days, months, or even years after being released from the parent plant, whereas most spores have shorter lifetimes. Also, unlike spores, seeds have a supply of stored food. Most seeds land close to their parent sporophyte plant, but some are carried long distances (up to hundreds of kilometers) by wind or animals. If conditions are favorable where it lands, the seed can emerge from dormancy and germinate, with its stored food providing critical support for growth as the sporophyte embryo emerges as a seedling. As we explore in the **Scientific Skills Exercise**, some seeds have germinated after more than 1,000 years.

CONCEPT CHECK 30.1

1. Contrast how sperm reach the eggs of seedless plants with how sperm reach the eggs of seed plants.
2. What features not present in seedless plants have contributed to the success of seed plants on land?
3. **WHAT IF?** ➤ If a seed could not enter dormancy, how might that affect the embryo's transport or survival?

For suggested answers, see Appendix A.

CONCEPT 30.2

Gymnosperms bear "naked" seeds, typically on cones

Extant seed plants form two sister clades: gymnosperms and angiosperms. Recall that gymnosperms have "naked" seeds exposed on sporophylls that usually form cones.

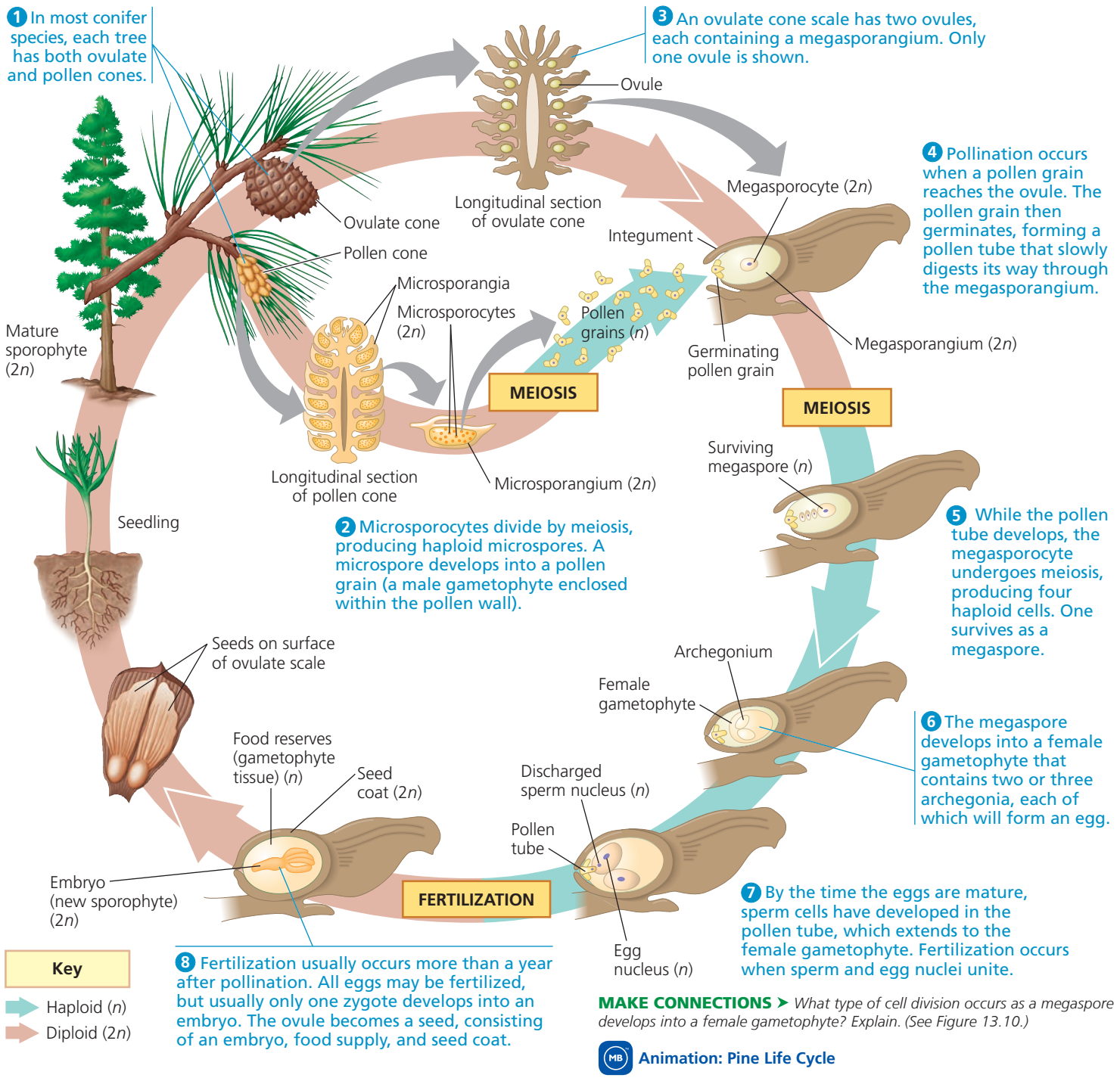
(Angiosperm seeds are enclosed in chambers that mature into fruits.) Most gymnosperms are cone-bearing plants called **conifers**, such as pines, firs, and redwoods.

The Life Cycle of a Pine

As you read earlier, seed plant evolution has included three key reproductive adaptations: the miniaturization of their gametophytes; the advent of the seed as a resistant, dispersible stage in the life cycle; and the appearance of pollen as an airborne agent that brings gametes together. **Figure 30.4** shows how these adaptations come into play during the life cycle of a pine, a familiar conifer.

The pine tree is the sporophyte; its sporangia are located on scalelike structures packed densely in cones. Like all seed plants, conifers are heterosporous. As such, they have two types of sporangia that produce two types of spores: microsporangia that produce microspores, and megasporangia that produce megaspores. In conifers, the two types of spores are produced by separate cones: small pollen cones and large ovulate cones.

Figure 30.4 The life cycle of a pine.



MAKE CONNECTIONS ➤ *What type of cell division occurs as a megaspore develops into a female gametophyte? Explain. (See Figure 13.10.)*

Animation: Pine Life Cycle

Pollen cones have a relatively simple structure: Their scales are modified leaves (microsporophylls) that bear microsporangia. Within each microsporangium, cells called microsporocytes undergo meiosis, producing haploid microspores. Each microspore develops into a pollen grain containing a male gametophyte. In conifers, the yellow pollen is released in large amounts and carried by the wind, dusting everything in its path.

Ovulate cones are more complex: their scales are compound structures composed of both modified leaves (megasporophylls bearing megasporangia) and modified stem tissue. Within the each megasporangium, megasporocytes undergo meiosis and produce haploid megaspores inside the ovule. Surviving megaspores develop into female gametophytes, which are retained within the sporangia.

In most pine species, each tree has both types of cones. From the time pollen and ovulate cones appear on the tree, it takes nearly three years for the male and female gametophytes to be produced and brought together and for mature seeds to form from fertilized ovules. The scales of each ovulate cone then separate, and seeds are dispersed by the wind. A seed that lands in a suitable environment germinates, its embryo emerging as a pine seedling.

Early Seed Plants and the Rise of Gymnosperms

The origins of characteristics found in pines and other living seed plants date back to the late Devonian period (380 million years ago). Fossils from that time reveal that some plants had acquired features that are also present in seed plants, such as megaspores and microspores. For example, *Archaeopteris* was a heterosporous tree with a woody stem. But it did not bear seeds and therefore is not classified as a seed plant. Growing up to 20 m tall, it had fernlike leaves.

The earliest evidence of seed plants comes from 360-million-year-old fossils of plants in the genus *Elkinsia* **(Figure 30.5)**. These and other early seed plants lived 55 million years before the first fossils classified as gymnosperms and more than 200 million years before the first fossils of angiosperms. These early seed plants became extinct, and we don't know which extinct lineage gave rise to the gymnosperms.

The oldest fossils of species from an extant lineage of gymnosperms are 305 million years old. These early gymnosperms lived in moist Carboniferous ecosystems that were dominated by lycophytes, horsetails, ferns, and other seedless vascular plants. As the Carboniferous period gave way to the Permian (299 to 252 million years ago), the climate became much drier. As a result, the lycophytes, horsetails, and ferns that dominated Carboniferous swamps were largely replaced by gymnosperms, which were better suited to the drier climate.

Figure 30.5 A fossil of the early seed plant *Elkinsia*.

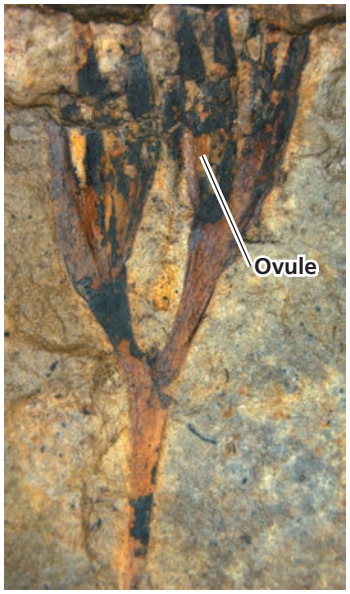
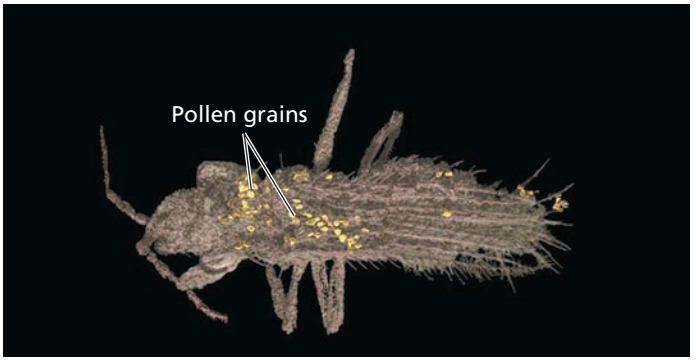


Figure 30.6 An ancient pollinator. This 110-million-year-old fossil shows pollen on an insect, the thrip *Gymnopollisthrips minor*. Structural features of the pollen suggest that it was produced by gymnosperms (most likely by species related to extant ginkgos or cycads). Although most gymnosperms today are wind-pollinated, many cycads are insect-pollinated.



Gymnosperms thrived as the climate dried, in part because they have the key terrestrial adaptations found in all seed plants, such as seeds and pollen. In addition, some gymnosperms were particularly well suited to arid conditions because of the thick cuticles and relatively small surface areas of their needle-shaped leaves.

Gymnosperms dominated terrestrial ecosystems throughout much of the Mesozoic era, which lasted from 252 to 66 million years ago. In addition to serving as the food supply for giant herbivorous dinosaurs, these gymnosperms were involved in many other interactions with animals. Recent fossil discoveries, for example, show that some gymnosperms were pollinated by insects more than 100 million years ago—the earliest evidence of insect pollination in any plant group **(Figure 30.6)**. Late in the Mesozoic, angiosperms began to replace gymnosperms in some ecosystems.

Gymnosperm Diversity

Although angiosperms now dominate most terrestrial ecosystems, gymnosperms remain an important part of Earth's flora. For example, vast regions in northern latitudes are covered by forests of conifers (see Figure 51.12).

Of the ten plant phyla (see Table 29.1), four are gymnosperms: Cycadophyta, Ginkgophyta, Gnetophyta, and Coniferophyta. It is uncertain how the four phyla of gymnosperms are related to each other. **Figure 30.7** surveys the diversity of extant gymnosperms.

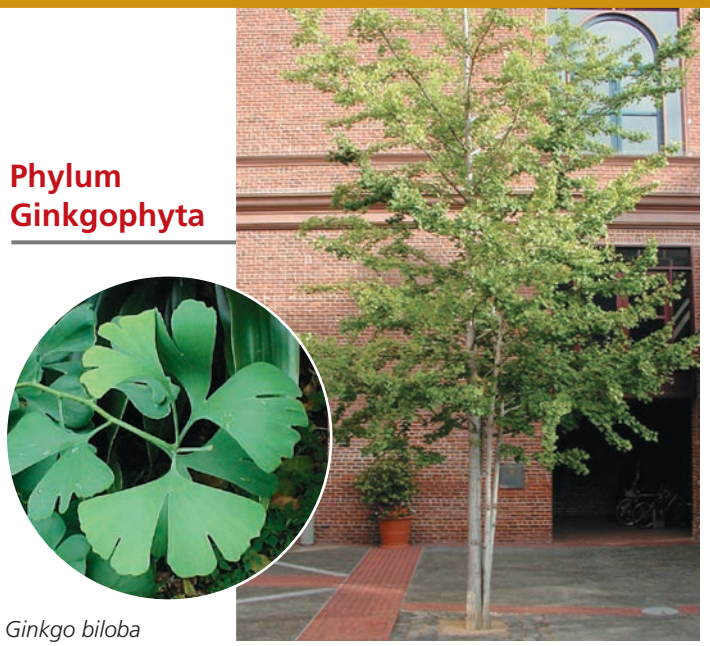
Figure 30.7 Exploring Gymnosperm Diversity

Phylum Cycadophyta

The 300 species of living cycads have large cones and palmlike leaves (true palm species are angiosperms). Unlike most seed plants, cycads have flagellated sperm, indicating their descent from seedless vascular plants that had motile sperm. Cycads thrived during the Mesozoic era, known as the age of cycads as well as the age of dinosaurs. Today, however, cycads are the most endangered of all plant groups: 75% of their species are threatened by habitat destruction and other human actions.

Encephalartos woodii

Phylum Ginkgophyta



Ginkgo biloba

Ginkgo biloba is the only surviving species of this phylum; like cycads, ginkgos have flagellated sperm. Also known as the maidenhair tree, *Ginkgo biloba* has deciduous fanlike leaves that turn gold in autumn. It is a popular ornamental tree in cities because it tolerates air pollution well. Landscapers often plant only pollen-producing trees because the fleshy seeds smell rancid as they decay.

Phylum Gnetophyta

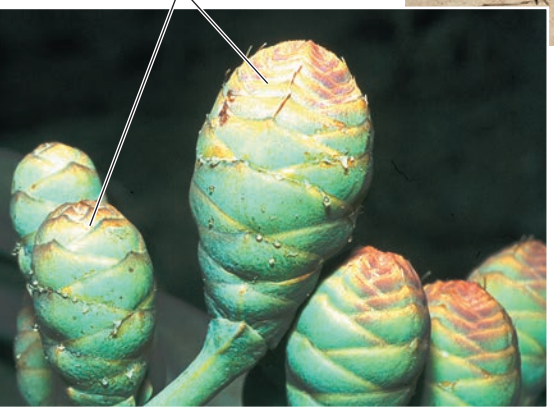
Phylum Gnetophyta includes plants in three genera: *Gnetum*, *Ephedra*, and *Welwitschia*. Some species are tropical, whereas others live in deserts. Although very different in appearance, the genera are grouped together based on molecular data.

Gnetum gnemon leaves

◀ ***Gnetum.*** This genus includes about 35 species of tropical trees, shrubs, and vines, mainly native to Africa and Asia. Their leaves look similar to those of flowering plants, and their seeds look somewhat like fruits.

▶ ***Welwitschia.*** This genus consists of one species, *Welwitschia mirabilis*, a plant that can live for thousands of years and is found only in the deserts of southwestern Africa. Its straplike leaves are among the largest leaves known.

Ovulate cones



▶ ***Ephedra.*** This genus includes about 40 species that inhabit arid regions worldwide. These desert shrubs, commonly called "Mormon tea", produce the compound ephedrine, which is used medicinally as a decongestant.

Phylum Coniferophyta

Phylum Coniferophyta, the largest gymnosperm phyla, consists of about 600 species of conifers (from the Latin *conus*, cone, and *ferre*, to carry), including many large trees. Most species have woody cones, but a few have fleshy cones. Some, such as pines, have needle-like leaves. Others, such as redwoods, have scale-like leaves. Some species dominate vast northern forests, whereas others are native to the Southern Hemisphere.

Most conifers are evergreens; they retain their leaves throughout the year. Even during winter, a limited amount of photosynthesis occurs on sunny days. When spring comes, conifers already have fully developed leaves that can take advantage of the sunnier, warmer days. Some conifers, such as the dawn redwood, tamarack, and larch, are deciduous trees that lose leaves each autumn.

▶ **Douglas fir.** This evergreen tree (*Pseudotduga menziesii*) provides more timber than any other North American tree species. Some uses include house framing, plywood, pulpwood for paper, railroad ties, and boxes and crates.

◀ **European larch.** The needle-like leaves of this deciduous conifer (*Larix decidua*) turn yellow before they are shed in autumn. Native to the mountains of central Europe, including Switzerland's Matterhorn, depicted here, this species is extremely cold-tolerant, able to survive winter temperatures that plunge to –50°C.

▶ **Common juniper.** The "berries" of the common juniper (*Juniperus communis*) are actually ovule-producing cones consisting of fleshy sporophylls.

◀ **Wollemi pine.** Survivors of a conifer group once known only from fossils, living *Wollemi* pines (*Wollemia nobilis*) were discovered in 1994 in a national park near Sydney, Australia. At that time, the species consisted of 40 known trees. As a result of conservation efforts, it is now widely propagated. The inset photo compares the leaves of this "living fossil" with actual fossils.

▶ **Sequoia.** This giant sequoia (*Sequoiadendron giganteum*) in California's Sequoia National Park weighs about 2,500 metric tons, equivalent to about 24 blue whales (the largest animals) or 40,000 people. The giant sequoia is one of the largest living organisms and also among the most ancient, with some individuals estimated to be between 1,800 and 2,700 years old. Their cousins, the coast redwoods (*Sequoia sempervirens*), grow to heights of more than 110 m (taller than the Statue of Liberty) and are found only in a narrow coastal strip of northern California and southern Oregon.

▶ **Bristlecone pine.** This species (*Pinus longaeva*), which is found in the White Mountains of California, includes some of the oldest living organisms, reaching ages of more than 4,600 years. One tree (not shown here) is called Methuselah because it may be the word's oldest living tree. To protect the tree, scientists keep its location a secret.

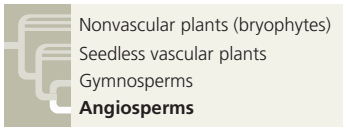
CONCEPT CHECK 30.2

1. Use examples from Figure 30.7 to describe how various gymnosperms are similar yet distinctive.
2. Explain how pollen cones differ from ovulate cones.
3. **MAKE CONNECTIONS ➤** Early seed plants in genus *Elkinsia* are a sister group to a clade consisting of gymnosperms and angiosperms. Draw a phylogenetic tree of seed plants that shows *Elkinsia*, gymnosperms, and angiosperms; date the branch points on this tree using fossil evidence. (See Figure 22.5.)

For suggested answers, see Appendix A.

CONCEPT 30.3

The reproductive adaptations of angiosperms include flowers and fruits



Commonly known as flowering plants, angiosperms are seed plants with the reproductive structures called flowers and fruits. The name *angiosperm* (from the Greek *angion*, container) refers to seeds contained in fruits. Angiosperms are the most diverse and widespread of all plants, with more than 250,000 species (about 90% of all plant species).

Characteristics of Angiosperms

All angiosperms are classified in a single phylum, Anthophyta. Before considering the evolution of angiosperms, we will examine two of their key adaptations—flowers and fruits—and the roles of these structures in the angiosperm life cycle.

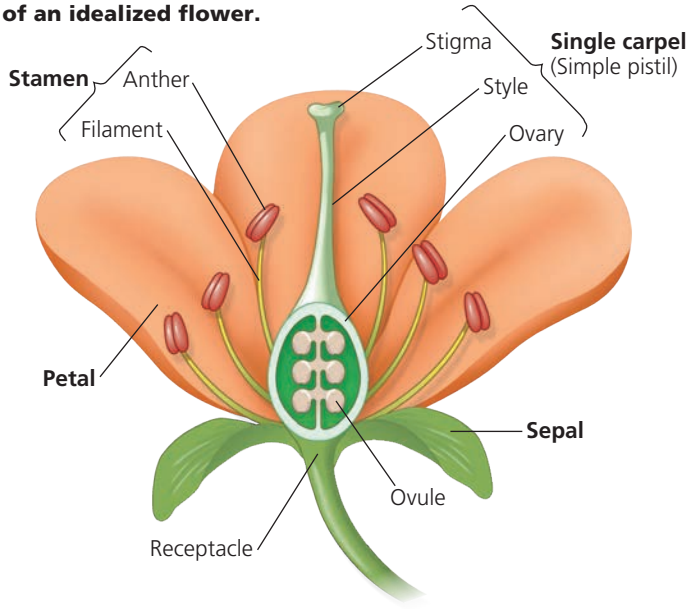
Flowers

The **flower** is a unique angiosperm structure that is specialized for sexual reproduction. In many angiosperm species, insects or other animals transfer pollen from one flower to the sex organs on another flower, which makes pollination more directed than the wind-dependent pollination of most species of gymnosperms. However, some angiosperms *are* wind-pollinated, particularly those species that occur in dense populations, such as grasses and tree species in temperate forests.

A flower is a specialized shoot that can have up to four types of modified leaves called floral organs: sepals, petals, stamens, and carpels **(Figure 30.8)**. Starting at the base of the flower are the **sepals**, which are usually green and enclose the flower before it opens (think of a rosebud). Interior to the sepals are the **petals**, which are brightly colored in most flowers and can aid in attracting pollinators. Flowers that are wind-pollinated, such as grasses, generally lack brightly colored parts. In all angiosperms, the sepals and petals are sterile floral organs, meaning that they do not produce sperm or eggs.

Within the petals are two types of fertile floral organs that produce spores, the stamens and carpels. Stamens and carpels

Figure 30.8 The structure of an idealized flower.

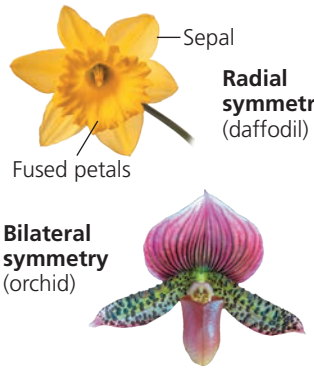


are sporophylls, modified leaves that are specialized for reproduction. **Stamens** are microsporophylls: They produce microspores that develop into pollen grains containing male gametophytes. A stamen consists of a stalk called the **filament** and a terminal sac, the **anther**, where pollen is produced. **Carpels** are megasporophylls: They produce megaspores that give rise to female gametophytes. The carpel is the "container" mentioned earlier in which seeds are enclosed; as such, it is a key structure that distinguishes angiosperms from gymnosperms. At the tip of the carpel is a sticky **stigma** that receives pollen. A **style** leads from the stigma to a structure at the base of the carpel, the **ovary**; the ovary contains one or more ovules. As in gymnosperms, each angiosperm ovule contains a female gametophyte. If fertilized, an ovule develops into a seed.

A flower may have one or more carpels. In many species, multiple carpels are fused into one structure. The term **pistil** is sometimes used to refer to a single carpel (a simple pistil) or two or more fused carpels (a compound pistil). Flowers also vary in symmetry **(Figure 30.9)** and other aspects of shape, as well as size, color, and odor. Much of this diversity results from adaptation to specific pollinators (see Figures 38.4 and 38.5).

Figure 30.9 Flower symmetry.

In radial symmetry, the sepals, petals, stamens, and carpels radiate out from a center. Any line through the central axis divides the flower into two equal parts. In bilateral symmetry, the flower can only be divided into two equal parts by a single line.



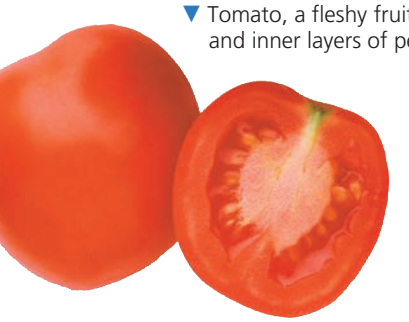
DRAW IT ➤ *Draw the single line that can divide the bilaterally symmetrical flower into two equal parts.*

Fruits

As seeds develop from ovules after fertilization, the ovary wall thickens and the ovary matures into a **fruit**. A pea pod is an example of a fruit, with seeds (mature ovules, the peas) encased in the ripened ovary (the pod).

Fruits protect seeds and aid in their dispersal. Mature fruits can be either fleshy or dry **(Figure 30.10)**. Tomatoes, plums, and grapes are examples of fleshy fruits, in which the wall (pericarp) of the ovary becomes soft during ripening. Dry fruits include beans, nuts, and grains. Some dry fruits split open at maturity to release seeds, whereas others remain closed. The dry, wind-dispersed fruits of grasses, harvested while on the plant, are major staple foods for humans. The cereal grains of maize, rice, wheat, and other grasses, though easily mistaken for seeds, are each actually a fruit with a dry outer covering (the former wall of the ovary) that adheres to the seed coat of the seed within.

Figure 30.10 Some variations in fruit structure.



▼ Tomato, a fleshy fruit with soft outer and inner layers of pericarp (fruit wall)

▼ Ruby grapefruit, a fleshy fruit with a firm outer layer and soft inner layer of pericarp

▼ Nectarine, a fleshy fruit with a soft outer layer and hard inner layer (pit) of pericarp

▼ Hazelnut, a dry fruit that remains closed at maturity

◀ Milkweed, a dry fruit that splits open at maturity

As shown in **Figure 30.11**, various adaptations of fruits and seeds help to disperse seeds (see also Figure 38.12). The seeds of some flowering plants, such as dandelions and maples, are contained within fruits that function like parachutes or propellers, adaptations that enhance dispersal by wind. Some fruits, such as coconuts, are adapted to dispersal by water. And the seeds of many angiosperms are carried by animals. Some angiosperms have fruits modified as burrs that cling to animal fur (or the clothes of humans). Others produce edible fruits, which are usually nutritious, sweet tasting, and vividly colored, advertising their ripeness. When an animal eats the fruit, it digests the fruit's fleshy part, but the tough seeds usually pass unharmed through the animal's

Figure 30.11 Fruit adaptations that enhance seed dispersal.

◀ Some plants have mechanisms that disperse seeds by explosive action.

▶ Wings enable maple fruits to be carried by the wind.

◀ Seeds within berries and other edible fruits are often dispersed in animal feces.

▶ The barbs of cockleburs facilitate seed dispersal by allowing the fruits to "hitchhike" on animals.

MB **Animation: Fruit Structure and Seed Dispersal**

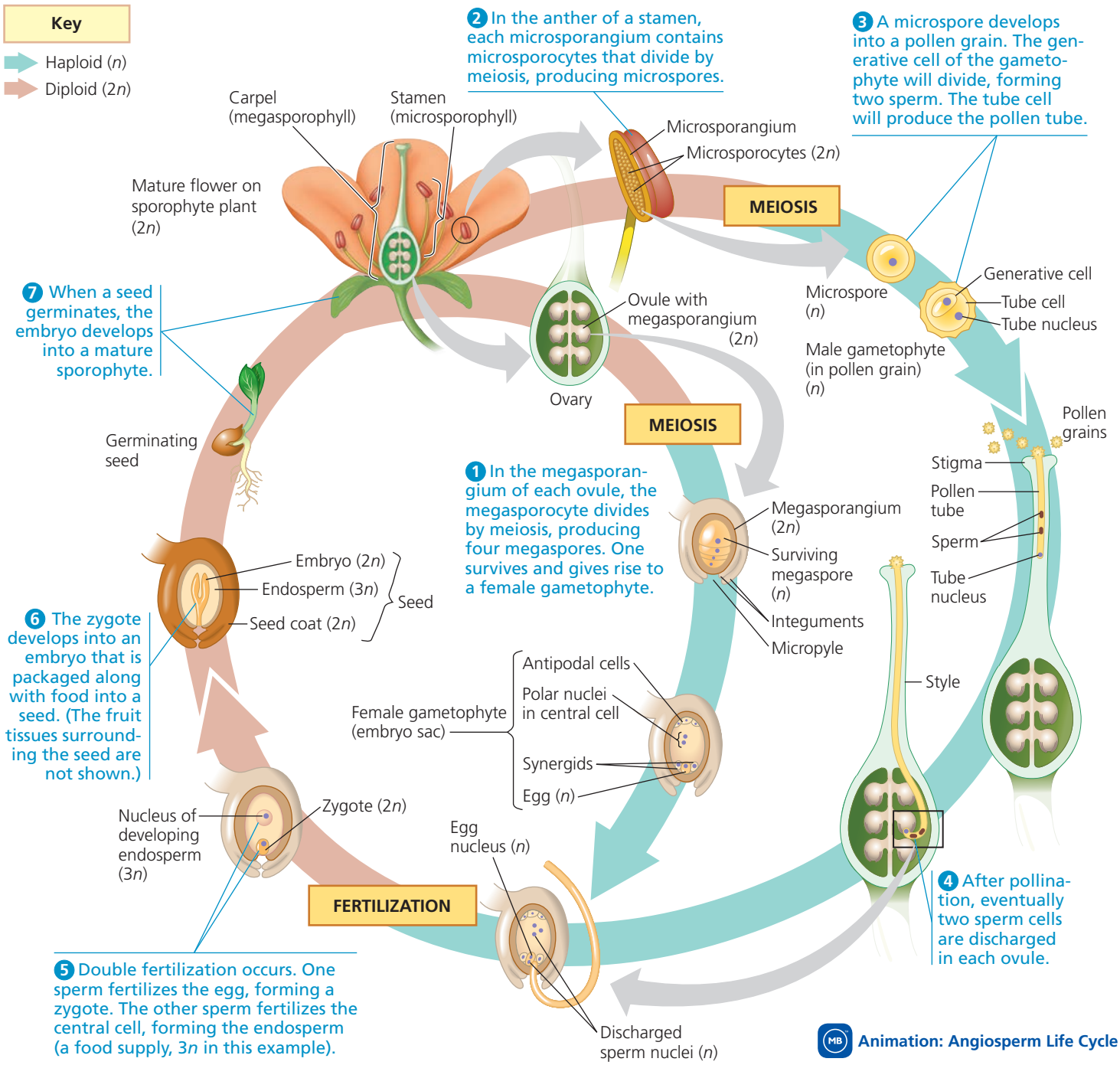
digestive tract. When the animal defecates, it may deposit the seeds, along with a supply of natural fertilizer, many kilometers from where the fruit was eaten.

The Angiosperm Life Cycle

You can follow a typical angiosperm life cycle in **Figure 30.12**. The flower of the sporophyte produces microspores that form male gametophytes and megaspores that form female gametophytes. The male gametophytes are in the pollen grains, which develop within microsporangia in the anthers. Each male gametophyte has two haploid cells: a *generative cell* that divides, forming two sperm, and a *tube cell* that produces a pollen tube. Each ovule, which develops in the ovary, contains a female gametophyte, also known as an **embryo sac**. The embryo sac consists of only a few cells, one of which is the egg.

After its release from the anther, the pollen is carried to the sticky stigma at the tip of a carpel. Although some flowers self-pollinate, most have mechanisms that ensure **cross-pollination**, which in angiosperms is the transfer of pollen from an anther of a flower on one plant to the stigma of a flower on another plant of the same species. Cross-pollination enhances genetic variability. In some species, stamens and

Figure 30.12 The life cycle of an angiosperm.



MB **Animation: Angiosperm Life Cycle**

carpels of a single flower may mature at different times, or they may be so arranged that self-pollination is unlikely.

The pollen grain absorbs water and germinates after it adheres to the stigma of a carpel. The tube cell produces a pollen tube that grows down within the style of the carpel. After reaching the ovary, the pollen tube penetrates through the **micropyle**, a pore in the integuments of the ovule, and discharges two sperm cells into the female gametophyte (embryo sac). One sperm fertilizes the egg, forming a diploid zygote. The other sperm fuses with the two nuclei in the large central cell of the female gametophyte, producing a triploid cell. This type of **double fertilization**, in which one fertilization event produces a zygote and the other produces a triploid cell, is unique to angiosperms.

After double fertilization, the ovule matures into a seed. The zygote develops into a sporophyte embryo with a rudimentary root and one or two seed leaves called **cotyledons**. The triploid central cell of the female gametophyte develops into **endosperm**, tissue rich in starch and other food reserves that nourish the developing embryo.

What is the function of double fertilization in angiosperms? One hypothesis is that double fertilization synchronizes the development of food storage in the seed with the development of the embryo. If a particular flower is not pollinated or sperm cells are not discharged into the embryo sac, fertilization does not occur, and neither endosperm nor embryo forms. So perhaps double fertilization is an adaptation that prevents flowering plants from squandering nutrients on infertile ovules.

Another type of double fertilization occurs in some gymnosperm species belonging to the phylum Gnetophyta. However, double fertilization in these species gives rise to two embryos rather than to an embryo and endosperm.

As you read earlier, the seed consists of the embryo, the endosperm, and a seed coat derived from the integuments. An ovary develops into a fruit as its ovules become seeds. After being dispersed, a seed may germinate if environmental conditions are favorable. The coat ruptures and the embryo emerges as a seedling, using food stored in the endosperm and cotyledons until it can produce its own food by photosynthesis.

Angiosperm Evolution

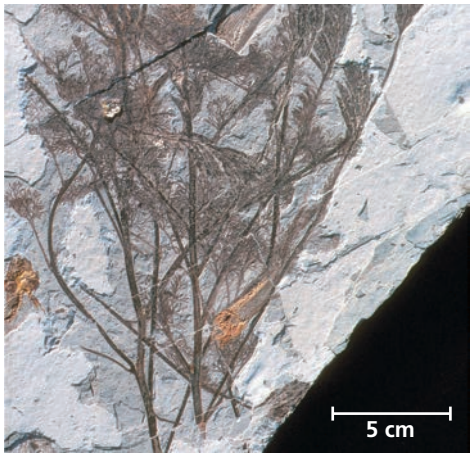
Charles Darwin once referred to the origin of angiosperms as an "abominable mystery." He was particularly troubled by the relatively sudden and geographically widespread appearance of angiosperms in the fossil record (about 100 million years ago, based on fossils known to Darwin). Recent fossil evidence and phylogenetic analyses have led to progress in solving Darwin's mystery, but we still do not fully understand how angiosperms arose from earlier seed plants.

Fossil Angiosperms

Angiosperms are now thought to have originated in the early Cretaceous period, about 140 million years ago. By the mid-Cretaceous (100 million years ago), angiosperms began to dominate some terrestrial ecosystems. Landscapes changed dramatically as conifers and other gymnosperms gave way to flowering plants in many parts of the world. The Cretaceous ended 66 million years ago with mass extinctions of dinosaurs and many other animal groups and with further increases in the diversity and importance of angiosperms.

What evidence suggests that angiosperms arose 140 million years ago? First, although pollen grains are common in rocks from the Jurassic period (201 to 145 million years ago), none of these pollen fossils have features characteristic of angiosperms, suggesting that angiosperms may have originated after the Jurassic. Indeed, the earliest fossils with distinctive angiosperm features are of 130-million-year-old pollen grains discovered in China, Israel, and England. Early fossils of larger flowering plant structures include those of *Archaefructus* **(Figure 30.13)** and *Leefructus*, both of which were discovered in China in rocks that are about 125 million years old. Overall, early angiosperm fossils indicate that the group arose and began to diversify over a 20- to 30-million-year period—a less

Figure 30.13 An early flowering plant.



(a) *Archaefructus sinensis*, a 125-million-year-old fossil. This herbaceous species had simple flowers and bulbous structures that may have served as floats, suggesting it was aquatic. Recent phylogenetic analyses indicate that *Archaefructus* may belong to the water lily group.

(b) Artist's reconstruction of *Archaefructus sinensis*

sudden event than was suggested by the fossils known during Darwin's lifetime.

Can we infer traits of the angiosperm common ancestor from traits found in early fossil angiosperms? *Archaefructus*, for example, was herbaceous and had bulbous structures that may have served as floats, suggesting it was aquatic. But investigating whether the angiosperm common ancestor was herbaceous and aquatic also requires examining fossils of other seed plants thought to have been closely related to angiosperms. All of those plants were woody, indicating that the common ancestor was probably woody and probably not aquatic. As we'll see, this conclusion has been supported by recent phylogenetic analyses.

Angiosperm Phylogeny

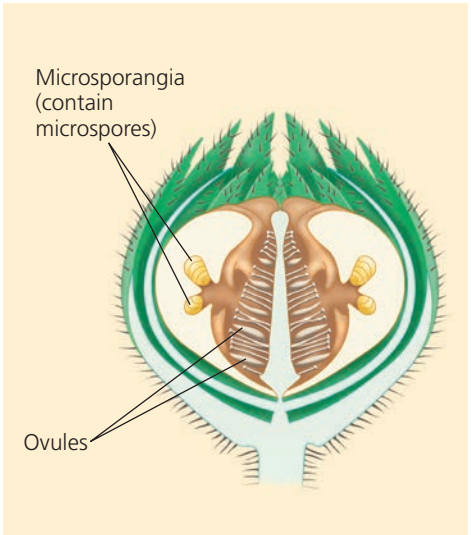
To shed light on the body plan of early angiosperms, scientists have sought to identify which seed plants, including fossil species, are most closely related to angiosperms. Molecular and morphological evidence suggests that extant gymnosperm lineages had diverged from the lineage leading to angiosperms by 305 million years ago. Note that this does not imply that angiosperms originated 305 million years ago, but that the most recent common ancestor of extant gymnosperms and angiosperms lived at that time. Indeed, angiosperms may be more closely related to several extinct lineages of woody seed plants than they are to gymnosperms. One such lineage is the Bennettitales, a group with flower-like structures that may have been pollinated by insects **(Figure 30.14a)**. However, the Bennettitales and other similar lineages of extinct woody seed plants did not have carpels or flowers and hence are not classified as angiosperms.

Making sense of the origin of angiosperms also depends on working out the order in which angiosperm clades diverged from one another. Here, dramatic progress has been made in recent years. Molecular and morphological evidence suggests that the shrub *Amborella trichopoda*, water lilies, and star anise are living representatives of lineages that diverged from other angiosperms early in the history of the group **(Figure 30.14b)**. *Amborella* is woody, supporting the conclusion mentioned earlier that the angiosperm common ancestor was probably woody. Like the Bennettitales, *Amborella*, water lilies, and star anise lack *vessel elements*, efficient water-conducting cells that are found in most present-day angiosperms. Overall, based on the features of ancestral species and angiosperms like *Amborella*, researchers have hypothesized that early angiosperms were woody shrubs that had small flowers and relatively simple water-conducting cells.

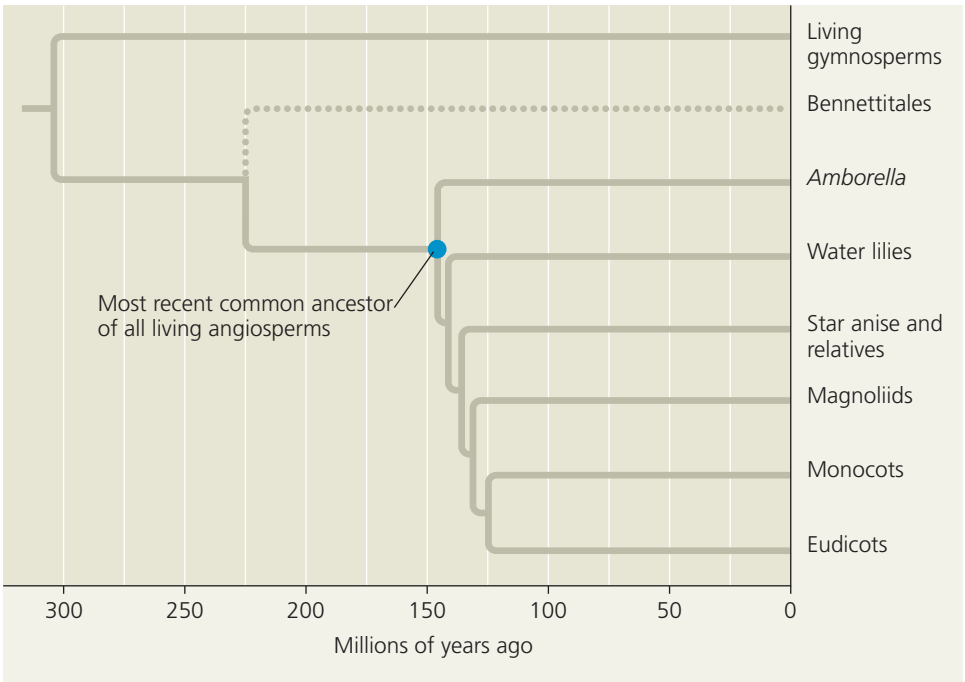
Evolutionary Links with Animals

Plants and animals have interacted for hundreds of millions of years, and those interactions have led to evolutionary change. For example, herbivores can reduce a plant's

Figure 30.14 Angiosperm evolutionary history.



(a) A close relative of the angiosperms? This reconstruction shows a longitudinal section through the flowerlike structures found in the Bennettitales, an extinct group of seed plants hypothesized to be more closely related to extant angiosperms than to extant gymnosperms.



(b) Angiosperm phylogeny. This tree represents a current hypothesis of angiosperm evolutionary relationships, based on morphological and molecular evidence. Angiosperms originated about 140 million years ago. The dotted line indicates the uncertain position of the Bennettitales, a possible sister group to extant angiosperms.

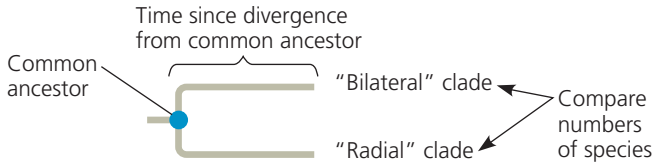
VISUAL SKILLS ➤ *Would the branching order of the phylogeny in (b) necessarily have to be redrawn if a 150-million-year-old fossil monocot were discovered? Explain.*

Figure 30.15 A bee pollinating a bilaterally symmetrical flower. To harvest nectar (a sugary solution secreted by flower glands) from this Scottish broom flower, a honeybee must land as shown. This releases a tripping mechanism that arches the flower's stamens over the bee and dusts it with pollen. Later, some of this pollen may rub off onto the stigma of the next flower of this species that the bee visits.

Video: Bee Pollinating

reproductive success by eating its roots, leaves, or seeds. As a result, if an effective defense against herbivores originates in a group of plants, those plants may be favored by natural selection—as will herbivores that overcome this new defense. Plant-pollinator and other mutually beneficial interactions also can have such reciprocal evolutionary effects.

Plant-pollinator interactions also may have affected the rates at which new species form. Consider the impact of a flower's symmetry (see Figure 30.9). On a flower with bilateral symmetry, an insect pollinator can obtain nectar only when approaching from a certain direction **(Figure 30.15)**. This constraint makes it more likely that pollen is placed on a part of the insect's body that will come into contact with the stigma of a flower of the same species. Such specificity of pollen transfer reduces gene flow between diverging populations and could lead to increased rates of speciation in plants with bilateral symmetry. This hypothesis can be tested using the approach illustrated in this diagram:



A key step in this approach is to identify cases in which a clade with bilaterally symmetric flowers shares an immediate common ancestor with a clade whose members have radially symmetric flowers. One recent study identified 19 pairs of closely related "bilateral" and "radial" clades. On average, the clade with bilaterally symmetric flowers had nearly 2,400 more species than did the related clade with radial symmetry. This result suggests that flower shape can affect the rate at which new species form, perhaps by affecting the behavior of insect pollinators. Overall, plant-pollinator interactions may have contributed to the increasing dominance of flowering plants in the Cretaceous period, helping to make angiosperms of central importance in ecological communities.

Angiosperm Diversity

From their humble beginnings in the Cretaceous period, angiosperms have diversified into more than 250,000 living species. Until the late 1990s, most systematists divided flowering plants into two groups, based partly on the number of cotyledons, or seed leaves, in the embryo. Species with one cotyledon were called **monocots**, and those with two were called **dicots**. Other features, such as flower and leaf structure, were also used to define the two groups. Recent DNA studies, however, indicate that the species traditionally called dicots are paraphyletic. The vast majority of species once categorized as dicots form a large clade, now known as **eudicots** ("true" dicots). **Figure 30.16** compares the main characteristics of monocots and eudicots. The rest of the former dicots are now grouped into four small lineages. Three of these lineages—*Amborella*, water lilies, and star anise and relatives—are informally called **basal angiosperms** because they diverged from other angiosperms early in the history of the group (see Figure 30.14b). A fourth lineage, the **magnoliids**, evolved later. **Figure 30.17** provides an overview of angiosperm diversity.

Figure 30.16 Characteristics of monocots and eudicots.

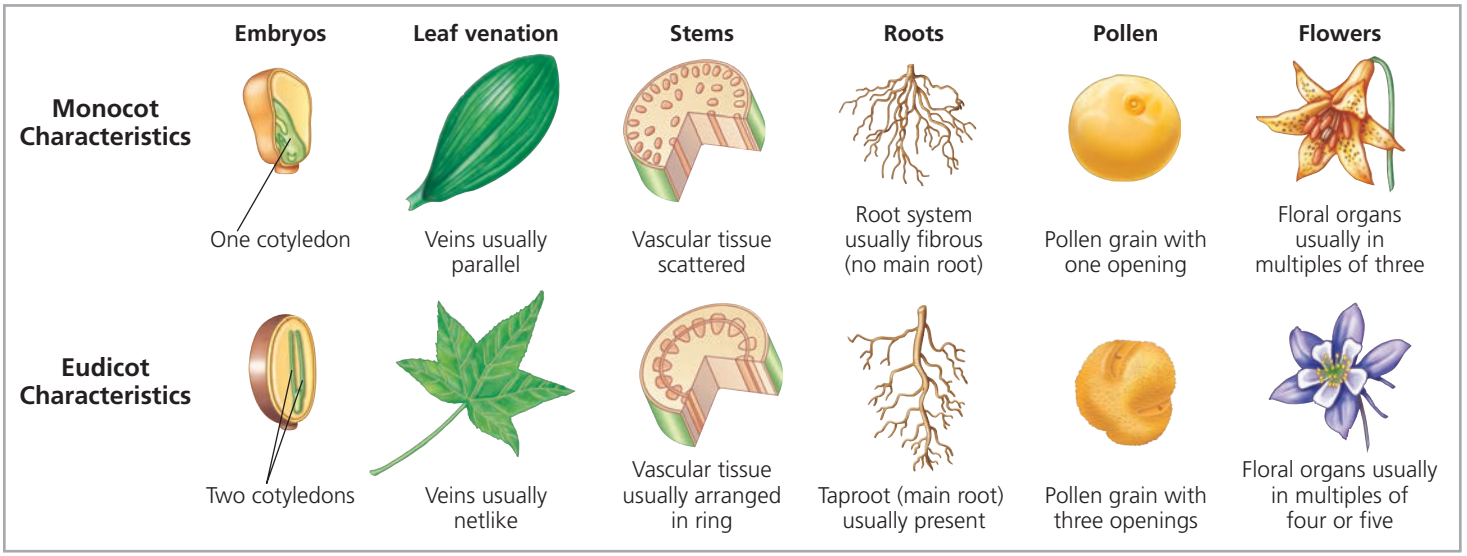


Figure 30.17 Exploring Angiosperm Diversity

Basal Angiosperms

Surviving basal angiosperms consist of three lineages comprising only about 100 species. The first lineage to have diverged from other angiosperms is represented today by a single species, *Amborella trichopoda* (right). The other surviving lineages diverged later: a clade that includes water lilies and a clade consisting of the star anise and its relatives.

◀ **Water lily (*Nymphaea "Rene Gerard"*).** Species of water lilies are found in aquatic habitats throughout the world. Water lilies belong to a clade that diverged from other angiosperms early in the group's history.

◀ **Star anise (*Illicium*).** This genus belongs to a third surviving lineage of basal angiosperms.

Amborella trichopoda. ▶ This small shrub, found only on the South Pacific island of New Caledonia, may be the sole survivor of a branch at the base of the angiosperm tree.

Magnoliids

Magnoliids consist of about 8,000 species, most notably magnolias, laurels, and black pepper plants. They include both woody and herbaceous species. Although they share some traits with basal angiosperms, such as a typically spiral rather than whorled arrangement of floral organs, magnoliids are more closely related to eudicots and monocots.

◀ **Southern magnolia (*Magnolia grandiflora*).** This member of the magnolia family is a large tree. The variety of southern magnolia shown here, called "Goliath," has flowers that measure up to about a foot across.

Monocots

About one-quarter of angiosperm species are monocots—about 70,000 species. Some of the largest groups are the orchids, grasses, and palms. Grasses include some of the most agriculturally important crops, such as maize, rice, and wheat.

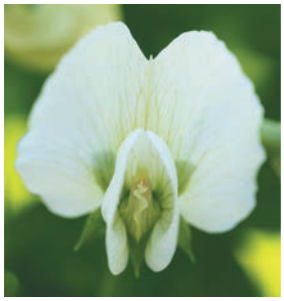
◀ **Orchid (*Lemboglossum rossii*)**

Barley (*Hordeum vulgare*), a grass ▶

▲ **Pygmy date palm (*Phoenix roebelenii*)**

Eudicots

More than two-thirds of angiosperm species are eudicots—roughly 170,000 species. The largest group is the legume family, which includes such crops as peas and beans. Also important economically is the rose family, which includes many plants with ornamental flowers as well as some species with edible fruits, such as strawberry plants and apple and pear trees. Most of the familiar flowering trees are eudicots, such as oak, walnut, maple, willow, and birch.



◀ **Snow pea (*Pisum sativum*), a legume**

Dog rose (*Rosa canina*), a wild rose ▶

◀ **Pyrenean oak (*Quercus pyrenaica*)**

CONCEPT CHECK 30.3

1. It is said that an oak is an acorn's way of making more acorns. Write an explanation that includes these terms: sporophyte, gametophyte, ovule, seed, ovary, and fruit.
2. Compare the process of double fertilization in angiosperms with that in gymnosperms.
3. **WHAT IF? ➤** Do speciation rates in closely related clades of flowering plants show that flower shape is *correlated with* the rate at which new species form or that flower shape is *responsible for* this rate? Explain.

For suggested answers, see Appendix A.

CONCEPT 30.4

Human welfare depends on seed plants

In forests and on farms, seed plants are key sources of food, fuel, wood products, and medicine. Our reliance on them makes the preservation of plant diversity critical.

Products from Seed Plants

Most of our food comes from angiosperms. Just six crops—maize, rice, wheat, potatoes, cassava, and sweet potatoes—yield 80% of all the calories consumed by humans. We also depend on angiosperms to feed livestock: It takes 5–7 kg of grain to produce 1 kg of grain-fed beef.

Today's crops are the products of artificial selection—the result of plant domestication that began about 12,000 years ago. To appreciate the scale of this transformation, note how the number and size of seeds in domesticated plants are greater than those of their wild relatives, as in the case of maize and the grass teosinte (see Figure 38.16). Scientists can glean information about domestication by comparing the genes of crops with those of wild relatives. With maize, dramatic changes such as increased cob size and loss of the hard coating around teosinte kernels may have been initiated by as few as five mutations.

Flowering plants also provide other edible products. Two popular beverages come from tea leaves and coffee beans, and you can thank the cacao tree for cocoa and chocolate. Spices are derived from various plant parts, such as flowers (cloves, saffron), fruits and seeds (vanilla, black pepper, mustard), leaves (basil, mint, sage), and even bark (cinnamon).

Many seed plants are sources of wood, which is absent in all living seedless plants. Wood consists of tough-walled xylem cells (see Figure 35.22). It is the primary source of fuel for much of the world, and wood pulp, typically derived from conifers such as fir and pine, is used to make paper. Wood remains the most widely used construction material.

For centuries, humans have also depended on seed plants for medicines. Many cultures use herbal remedies, and scientists have extracted and identified medicinally active compounds from many of these plants and later synthesized them. Willow leaves and bark have long been used in pain-relieving remedies, including prescriptions by the Greek physician Hippocrates. In the 1800s, scientists traced the willow's medicinal property to the chemical salicin. A synthesized derivative, acetylsalicylic acid, is what we call aspirin. Plants are also a direct source of medicinal compounds **(Table 30.1)**. In the United States, about 25% of prescription drugs contain an active ingredient from plants, usually seed plants.

Table 30.1 Examples of Plant-Derived Medicines

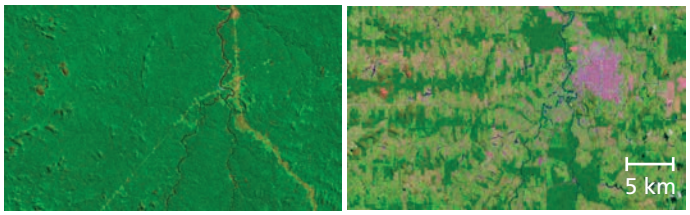
Compound	Source	Use
Atropine	Belladonna plant	Eye pupil dilator
Digitalin	Foxglove	Heart medication
Menthol	Eucalyptus tree	Throat soother
Quinine	Cinchona tree	Malaria preventive
Taxol	Pacific yew	Ovarian cancer drug
Tubocurarine	Curare tree	Muscle relaxant
Vinblastine	Periwinkle	Leukemia drug

Threats to Plant Diversity

Although plants may be a renewable resource, plant diversity is not. The exploding human population and its demand for space and resources are threatening plant species across the globe. The problem is especially severe in the tropics, where more than two-thirds of the human population live and where population growth is fastest. About 63,000 km^2 (15 million acres) of tropical rain forest are cleared each year **(Figure 30.18)**, a rate that would completely eliminate the remaining 11 million km^2 of tropical forests in 175 years. The loss of forests reduces the absorption of atmospheric carbon dioxide (CO_2) that occurs during photosynthesis, potentially contributing to global warming. Also, as forests disappear, so do large numbers of plant species. Of course, once a species becomes extinct, it can never return.

The loss of plant species is often accompanied by the loss of insects and other rain forest animals. Scientists estimate that if current rates of loss in the tropics and elsewhere

Figure 30.18 Clear-cutting of tropical forests. Over the past several hundred years, nearly half of Earth's tropical forests have been cut down and converted to farmland and other uses. A satellite image from 1975 (left) shows a dense forest in Brazil. By 2012, much of this forest had been cut down. Deforested and urban areas are shown as light purple.



continue, 50% or more of Earth's species will become extinct within the next few centuries. Such losses would constitute a global mass extinction, rivaling the Permian and Cretaceous mass extinctions and forever changing the evolutionary history of plants (and many other organisms).

Many people have ethical concerns about contributing to the extinction of species. In addition, there are practical reasons to be concerned about the loss of plant diversity. So far, we have explored the potential uses of only a tiny fraction of the more than 290,000 known plant species. For example, almost all our food is based on the cultivation of only about two dozen species of seed plants. And fewer than 5,000 plant species have been studied as potential sources of medicines. The tropical rain forest may be a medicine chest of healing plants that could be extinct before we even know they exist. If we begin to view rain forests and other ecosystems as living treasures that can regenerate only slowly, we may learn to harvest their products at sustainable rates.

CONCEPT CHECK 30.4

1. Explain why plant diversity can be considered a nonrenewable resource.
2. **WHAT IF?** ➤ How could phylogenies be used to help researchers search more efficiently for novel medicines derived from seed plants?

For suggested answers, see Appendix A.

30 Chapter Review

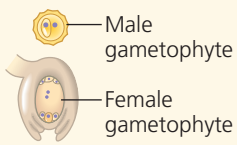
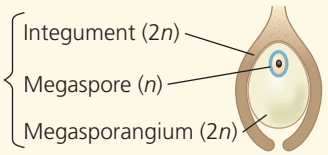
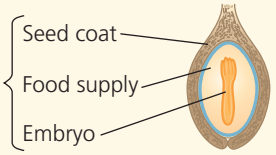
Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 30.1

Seeds and pollen grains are key adaptations for life on land *(pp. 689–691)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

Five Derived Traits of Seed Plants	
Reduced gametophytes	Microscopic male and female gametophytes (*n*) are nourished and protected by the sporophyte (2*n*) 
Heterospory	Microspore (gives rise to a male gametophyte)  Megaspore (gives rise to a female gametophyte) 
Ovules	Ovule (gymnosperm) 
Pollen	Pollen grains make water unnecessary for fertilization 
Seeds	Seeds: survive better than unprotected spores, can be transported long distances 

? *Describe how the parts of an ovule (integument, megaspore, megasporangium) correspond to the parts of a seed.*

CONCEPT 30.2

Gymnosperms bear "naked" seeds, typically on cones *(pp. 691–696)*

- Dominance of the sporophyte generation, the development of seeds from fertilized ovules, and the role of pollen in transferring sperm to ovules are key features of a typical gymnosperm life cycle.
- Gymnosperms appear early in the plant fossil record and dominated many Mesozoic terrestrial ecosystems. Living seed plants can be divided into two monophyletic groups: gymnosperms and angiosperms. Extant gymnosperms include cycads, *Ginkgo biloba*, gnetophytes, and **conifers**.

? *Although there are fewer than 1,000 species of gymnosperms, the group is still very successful in terms of its evolutionary longevity, adaptations, and geographic distribution. Explain.*

CONCEPT 30.3

The reproductive adaptations of angiosperms include flowers and fruits *(pp. 696–703)*

- **Flowers** generally consist of four types of modified leaves: **sepals, petals, stamens** (which produce pollen), and **carpels** (which produce ovules). **Ovaries** ripen into **fruits**, which often carry seeds by wind, water, or animals to new locations.
- Flowering plants originated about 140 million years ago, and by the mid-Cretaceous (100 mya) had begun to dominate some terrestrial ecosystems. Fossils and phylogenetic analyses offer insights into the origin of flowers.
- Several groups of **basal angiosperms** have been identified. Other major clades of angiosperms include **magnoliids, monocots**, and **eudicots**.
- Pollination and other interactions between angiosperms and animals may have contributed to the success of flowering plants during the last 100 million years.

? *Explain why Darwin called the origin of angiosperms an "abominable mystery," and describe what has been learned from fossil evidence and phylogenetic analyses.*

CONCEPT 30.4

Human welfare depends on seed plants

(pp. 703–704)

- Humans depend on seed plants for products such as food, wood, and many medicines.
- Destruction of habitat threatens the extinction of many plant species and the animal species they support.

? *Explain why destroying the remaining tropical forests might harm humans and lead to a mass extinction.*

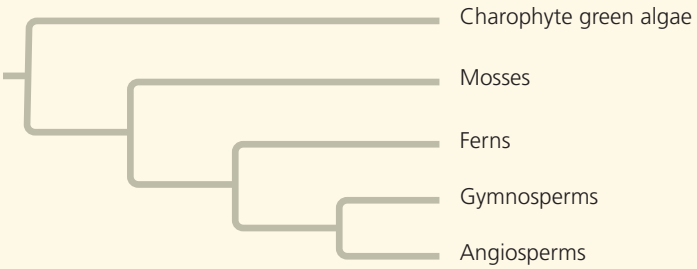
TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–5 can be found in the Study Area in MasteringBiology.

6. DRAW IT Use the letters a–d to label where on the phylogenetic tree each of the following derived characters appears.

(A) flowers
(B) embryos
(C) seeds
(D) vascular tissue

PRACTICE TEST
goo.gl/iAsVgL



7. EVOLUTION CONNECTION The history of life has been punctuated by several mass extinctions. For example, the impact of a meteorite may have wiped out most of the dinosaurs and many forms of marine life at the end of the Cretaceous period (see Concept 25.4). Fossils indicate that plants were less severely affected by this mass extinction. What adaptations may have enabled plants to withstand this disaster better than animals?

8. SCIENTIFIC INQUIRY • DRAW IT As will be described in detail in Concept 38.1, the female gametophyte of angiosperms typically has seven cells, one of which, the central cell, contains two haploid nuclei. After double fertilization, the central cell develops into endosperm, which is triploid. Because magnoliids, monocots, and eudicots typically have female gametophytes with seven cells and triploid endosperm, scientists assumed that this was the ancestral state for angiosperms. Consider, however, the following recent discoveries:

- Our understanding of angiosperm phylogeny has changed to that shown in Figure 30.14b.
- *Amborella trichopoda* has eight-celled female gametophytes and triploid endosperm.
- Water lilies and star anise have four-celled female gametophytes and diploid endosperm.

(a) Draw a phylogeny of the angiosperms (see Figure 30.14b), incorporating the data given above about the number of cells in female gametophytes and the ploidy of the endosperm. Assume that all of the star anise relatives have four-celled female gametophytes and diploid endosperm.
(b) What does your labeled phylogeny suggest about the evolution of the female gametophyte and endosperm in angiosperms?

9. WRITE ABOUT A THEME: ORGANIZATION Cells are the basic units of structure and function in all organisms. A key feature in the life cycle of plants is the alternation of multicellular haploid and diploid generations. Imagine a lineage of flowering plants in which mitotic cell division did not occur between the events of meiosis and fertilization (see Figure 30.12). In a short essay (100–150 words), describe how this change in the timing of cell division would affect the structure and life cycle of plants in this lineage.

10. SYNTHESIZE YOUR KNOWLEDGE

This photograph shows a dandelion seed in flight. Describe how seeds and other adaptations in seed plants contributed to the rise of seed plants and their dominant role in plant communities today.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

31 Introduction to Fungi

Figure 31.1 What role does this mushroom play in the forest?

KEY CONCEPTS

31.1 Fungi are heterotrophs that feed by absorption

31.2 Fungi produce spores through sexual or asexual life cycles

31.3 The ancestor of fungi was an aquatic, single-celled, flagellated protist

31.4 Fungi have radiated into a diverse set of lineages

31.5 Fungi play key roles in nutrient cycling, ecological interactions, and human welfare

▼ *Cortinarius caperatus*, another species that may transfer sugars between trees

Hidden Networks

Walking through a pine forest in Switzerland, you might spot some small reddish mushrooms in the genus *Russula* scattered here and there beneath the towering trees **(Figure 31.1)**. These little mushrooms are just the aboveground portion of a vast network of filaments located beneath the forest floor. As they grow, these fungal filaments absorb nutrients, some of which they transfer to the roots of trees. In turn, the trees provide the fungi with sugars produced in photosynthesis. A 2016 study found that these fungal filaments can even transfer sugars between the trees of different species. Thus, sugars produced by one tree might actually nourish the cells of nearby trees—adding a new level of complexity to our understanding of life in a forest.

The hidden network of *Russula* filaments is a fitting symbol of the neglected grandeur of the kingdom Fungi. Most of us are barely aware of these eukaryotes beyond the mushrooms we eat or the occasional brush with athlete's foot. Yet fungi are a huge and important component of the biosphere. While about 100,000 species have been described, there may be as many as 1.5 million species of fungi. Some fungi are exclusively single-celled, though most have complex multicellular bodies. These diverse organisms are found in just about every imaginable terrestrial and aquatic habitat.

Fungi are not only diverse and widespread but also essential for the well-being of most ecosystems. They break down organic material and recycle nutrients, allowing other organisms to assimilate essential chemical elements. Humans make use of fungi as a food source, for applications in agriculture and forestry, and in

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

manufacturing products ranging from bread to antibiotics. But it is also true that some fungi cause disease in plants and animals.

In this chapter, we will investigate the structure and evolutionary history of fungi, survey the major groups of fungi, and discuss their ecological and commercial significance.

CONCEPT 31.1

Fungi are heterotrophs that feed by absorption

Despite their vast diversity, all fungi share some key traits: most importantly, the way they derive nutrition. In addition, many fungi grow by forming multicellular filaments, a body structure that plays an important role in how they obtain food.

Nutrition and Ecology

Like animals, fungi are heterotrophs: They cannot make their own food as plants and algae can. But unlike animals, fungi do not ingest (eat) their food. Instead, a fungus absorbs nutrients from the environment outside of its body. Many fungi do this by secreting hydrolytic enzymes into their surroundings. These enzymes break down complex molecules to smaller organic compounds that the fungi can absorb into their cells and use. Other fungi use enzymes to penetrate the walls of cells, enabling the fungi to absorb nutrients from the cells. Collectively, the different enzymes found in various fungal species can digest compounds from a wide range of sources, living or dead.

This diversity of food sources corresponds to the varied roles of fungi in ecological communities: Different species live as decomposers, parasites, or mutualists. Fungi that are decomposers break down and absorb nutrients from nonliving organic material, such as fallen logs, animal corpses, and the wastes of living organisms. Parasitic fungi absorb nutrients from the cells of living hosts. Some parasitic fungi are pathogenic, including many species that cause diseases in plants and others that cause diseases in animals. Mutualistic fungi also absorb nutrients from a host organism, but they reciprocate with actions that benefit the host. For example, mutualistic fungi that live within the digestive tracts of certain termite species use their enzymes to break down wood, as do mutualistic protists in other termite species (see Figure 28.27).

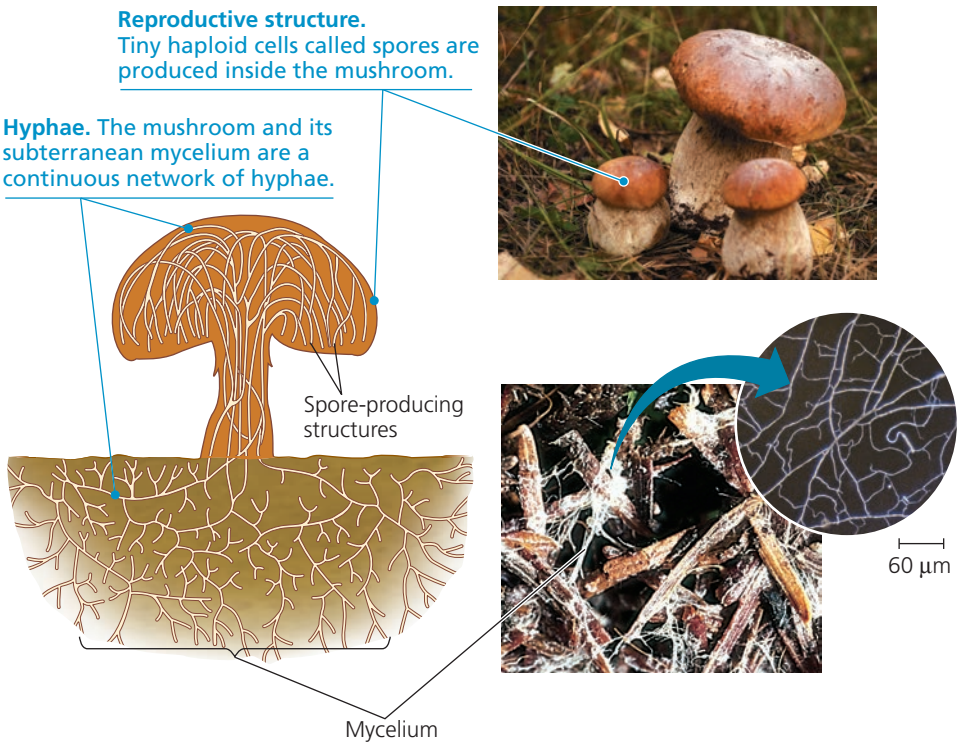
The versatile enzymes that enable fungi to digest a wide range of food sources are not the only reason for their ecological success. Another important factor is how their body structure increases the efficiency of nutrient absorption.

Body Structure

The most common fungal body structures are multicellular filaments and single cells (**yeasts**). Many fungal species can grow as both filaments and yeasts, but even more grow only as filaments; relatively few species grow only as single-celled yeasts. Yeasts often inhabit moist environments, including plant sap and animal tissues, where there is a ready supply of soluble nutrients, such as sugars and amino acids.

The morphology of multicellular fungi enhances their ability to grow into and absorb nutrients from their surroundings **(Figure 31.2)**. The bodies of these fungi typically form a network of tiny filaments called **hyphae** (singular, *hypha*). Hyphae consist of tubular cell walls surrounding the plasma membrane and cytoplasm of the cells. The cell walls are strengthened by **chitin**, a strong but flexible polysaccharide. Chitin-rich walls can enhance feeding by absorption. As a fungus absorbs nutrients from its environment, the concentrations of those nutrients in its cells increases, causing water to move into the cells by osmosis. The movement of water

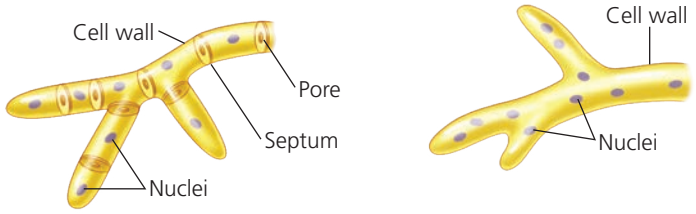
Figure 31.2 Structure of a multicellular fungus. The top photograph shows the sexual structures, in this case called mushrooms, of the penny bun fungus (*Boletus edulis*). The bottom photograph shows a mycelium growing on fallen conifer needles. The inset SEM shows hyphae.



? *Although the mushrooms in the top photograph appear to be different individuals, could their DNA be identical? Explain.*

MB **Animation: Fungal Growth and Nutrition**

Figure 31.3 Two forms of hyphae.



(a) Septate hypha (b) Coenocytic hypha

into fungal cells creates pressure that could cause them to burst if they were not surrounded by a rigid cell wall.

Another important structural feature of most fungi is that their hyphae are divided into cells by cross-walls, or **septa** (singular, *septum*) **(Figure 31.3a)**. Septa generally have pores large enough to allow ribosomes, mitochondria, and even nuclei to flow from cell to cell. Some fungi lack septa **(Figure 31.3b)**. Known as **coenocytic fungi**, these organisms consist of a continuous cytoplasmic mass having hundreds or thousands of nuclei. As we'll describe later, the coenocytic condition results from the repeated division of nuclei without cytokinesis.

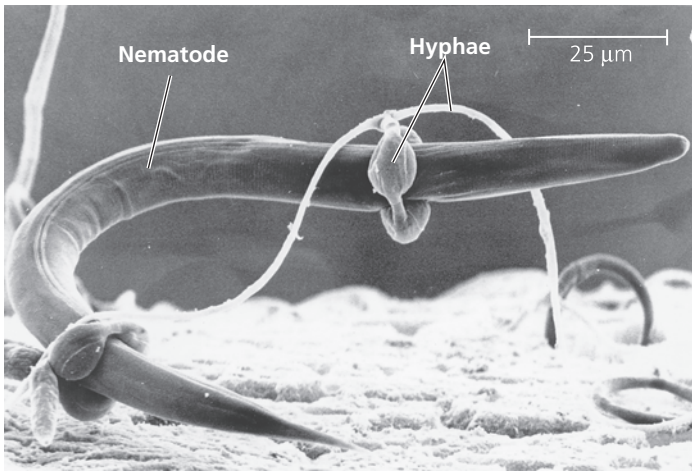
Fungal hyphae form an interwoven mass called a **mycelium** (plural, *mycelia*) that infiltrates the material on which the fungus feeds (see Figure 31.2). The structure of a mycelium maximizes its surface-to-volume ratio, making feeding very efficient. Just 1 cm^3 of rich soil may contain as much as 1 km of hyphae with a total surface area of 300 cm^2 in contact with the soil. A fungal mycelium grows rapidly, as proteins and other materials synthesized by the fungus move through cytoplasmic streaming to the tips of the extending hyphae. The fungus concentrates its energy and resources on adding hyphal length and thus overall absorptive surface area, rather than on increasing hyphal girth. Multicellular fungi are not motile in the typical sense—they cannot run, swim, or fly in search of food or mates. However, as they grow, such fungi can move into new territory, swiftly extending the tips of their hyphae.

Specialized Hyphae in Mycorrhizal Fungi

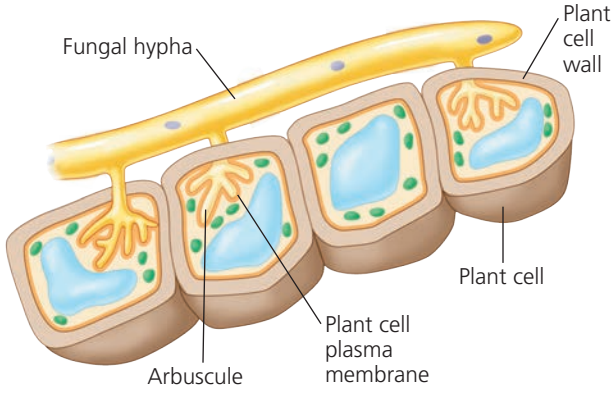
Some fungi have specialized hyphae that allow them to feed on living animals **(Figure 31.4a)**, while others have modified hyphae called *haustoria* that enable them to extract nutrients from plants. Our focus here, however, will be on fungi that have specialized branching hyphae such as **arbuscules (Figure 31.4b)** that they use to exchange nutrients with their plant hosts. Such mutually beneficial relationships between fungi and plant roots are called **mycorrhizae** (the term means "fungus roots").

Mycorrhizal fungi (fungi that form mycorrhizae) can improve delivery of phosphate ions and other minerals to plants because the vast mycelial networks of the fungi are more efficient than the plants' roots at acquiring these minerals from the soil. In exchange, the plants supply the fungi with organic nutrients such as carbohydrates.

Figure 31.4 Specialized hyphae.



(a) Hyphae adapted for trapping and killing prey. In *Arthrobotrys*, a soil fungus, portions of the hyphae are modified as hoops that can constrict around a nematode (roundworm) in less than a second. The growing hyphae then penetrate the worm's body, and the fungus digests its prey's inner tissues (SEM).



(b) Arbuscules. Some mutualistic fungi have specialized hyphae called arbuscules that can exchange nutrients with living plant cells. Arbuscules remain separated from a plant cell's cytoplasm by the plasma membrane of the plant cell (orange).

There are two main types of mycorrhizal fungi (see Figure 37.15). **Ectomycorrhizal fungi** (from the Greek *ektos*, out) form sheaths of hyphae over the surface of a root and typically grow into the extracellular spaces of the root cortex. **Arbuscular mycorrhizal fungi** extend arbuscules through the root cell wall and into tubes formed by invagination (pushing inward, as in Figure 31.4b) of the root cell plasma membrane. In the **Scientific Skills Exercise**, you'll compare genomic data from fungi that form mycorrhizae and fungi that do not.

Mycorrhizae are enormously important both in natural ecosystems and in agriculture. Almost all vascular plants have mycorrhizae and rely on their fungal partners for essential nutrients. Many studies have shown the significance of mycorrhizae by comparing the growth of plants with and without them. Foresters commonly inoculate pine seedlings with mycorrhizal fungi to promote growth. In the absence of human intervention, mycorrhizal fungi colonize

SCIENTIFIC SKILLS EXERCISE

Interpreting Genomic Data and Generating Hypotheses

What Can Genomic Analysis of a Mycorrhizal Fungus Reveal About Mycorrhizal Interactions? The first genome of a mycorrhizal fungus to be sequenced was that of the basidiomycete *Laccaria bicolor* (see photo). In nature, *L. bicolor* is a common ectomycorrhizal fungus of trees such as poplar and fir, as well as a free-living soil organism. In forest nurseries, it is often added to soil to enhance seedling growth. The fungus can easily be grown alone in culture and can establish mycorrhizae with tree roots in the laboratory. Researchers hope that studying the genome of *Laccaria* will yield clues to the processes by which it interacts with its mycorrhizal partners—and by extension, to mycorrhizal interactions involving other fungi.

How the Study Was Done Using the whole-genome shotgun method (see Figure 20.2) and bioinformatics, researchers sequenced the genome of *L. bicolor* and compared it with the genomes of some nonmycorrhizal basidiomycete fungi. The team used microarrays to compare gene expression levels for different protein-coding genes and for the same genes in a mycorrhizal mycelium and a free-living mycelium. They could thus identify the genes for fungal proteins that are made specifically in mycorrhizae.

Data from the Study

Table 1 Numbers of Genes in *L. bicolor* and Four Nonmycorrhizal Fungal Species

	L. bicolor	1	2	3	4
Protein-coding genes	20,614	13,544	10,048	7,302	6,522
Genes for membrane transporters	505	412	471	457	386
Genes for small secreted proteins (SSPs)	2,191	838	163	313	58

Table 2 *L. bicolor* Genes Most Highly Upregulated in Ectomycorrhizal Mycelium (ECM) of Douglas Fir or Poplar vs. Free-Living Mycelium (FLM)

Protein ID	Protein Feature or Function	Douglas Fir ECM/FLM Ratio	Poplar ECM/FLM Ratio
298599	SSP	22,877	12,913
293826	Enzyme inhibitor	14,750	17,069
333839	SSP	7,844	1,931
316764	Enzyme	2,760	1,478

Data from F. Martin et al., The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis, *Nature* 452:88–93 (2008).

INTERPRET THE DATA

1. (a) In Table 1, which fungal species has the most genes encoding membrane transporters (membrane transport proteins; see Concept 8.2)? (b) Why might these genes be of particular importance to *L. bicolor*?
2. The phrase "small secreted proteins" (SSPs) refers to proteins less than 100 amino acids in length that the fungi secrete; their function is not yet known. (a) Describe the Table 1 data on SSPs. (b) The researchers found that the SSP genes shared a common feature that indicated the encoded proteins were destined for secretion. Based on Figure 17.22 and the text discussion of that figure, predict what this common feature of the SSP genes was. (c) Suggest a hypothesis for the roles of SSPs in mycorrhizae.
3. Table 2 shows data from gene expression studies for the four *L. bicolor* genes whose transcription was most increased ("upregulated") in mycorrhizae. (a) For the gene encoding the first protein listed, what does the number 22,877 indicate? (b) Do the data in Table 2 support your hypothesis in 2(c)? Explain. (c) Compare the data for poplar mycorrhizae with those for Douglas fir and hypothesize what might account for any differences.

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

soils by dispersing haploid cells called **spores** that form new mycelia after germinating. Spore dispersal is a key component of how fungi reproduce and spread to new areas, as we discuss next.

CONCEPT CHECK 31.1

1. How do mycorrhizal fungi establish mutually beneficial relationships with host plants?
2. **WHAT IF?** ➤ Suppose a certain fungus is a mutualist that lives within an insect host, yet its ancestors were parasites that grew in and on the insect's body. What derived traits might you find in this mutualistic fungus?
3. **MAKE CONNECTIONS** ➤ Review Figure 11.4 and Figure 11.6. If a plant has mycorrhizae, where might carbon that enters the plant's stomata as CO_2 eventually be deposited: in the plant, in the fungus, or both? Explain.

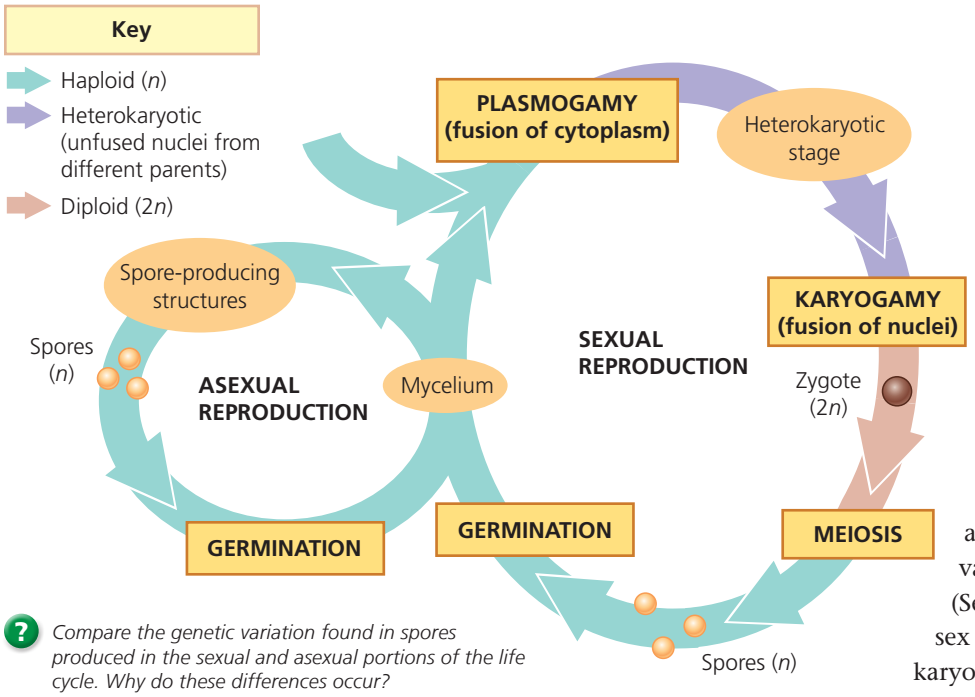
For suggested answers, see Appendix A.

CONCEPT 31.2

Fungi produce spores through sexual or asexual life cycles

Most fungi propagate themselves by producing vast numbers of spores, either sexually or asexually. For example, puffballs, the reproductive structures of certain fungal species, may release trillions of spores (see Figure 31.17). Spores can be carried long distances by wind or water. If they land in a moist place where there is food, they germinate, producing a new mycelium. To appreciate how effective spores are at dispersing, leave a slice of melon exposed to the air. Even without a visible source of spores nearby, within a week, you will likely observe fuzzy mycelia growing from microscopic spores that have fallen onto the melon.

Figure 31.5 Generalized life cycle of fungi. Many fungi reproduce both sexually and asexually, as shown here; others, however, reproduce only sexually or asexually.



? *Compare the genetic variation found in spores produced in the sexual and asexual portions of the life cycle. Why do these differences occur?*

Figure 31.5 generalizes the many different life cycles that can produce fungal spores. In this section, we will survey the main aspects of sexual and asexual reproduction in fungi.

Sexual Reproduction

The nuclei of fungal hyphae and the spores of most fungi are haploid, although many species have transient diploid stages that form during sexual life cycles. Sexual reproduction often begins when hyphae from two mycelia release sexual signaling molecules called **pheromones**. If the mycelia are of different mating types, the pheromones from each partner bind to receptors on the other, and the hyphae extend toward the source of the pheromones. When the hyphae meet, they fuse. In species with such a "compatibility test," this process contributes to genetic variation by preventing hyphae from fusing with other hyphae from the same mycelium or another genetically identical mycelium.

The union of the cytoplasms of two parent mycelia is known as **plasmogamy** (see Figure 31.5). In most fungi, the haploid nuclei contributed by each parent do not fuse right away. Instead, parts of the fused mycelium contain coexisting, genetically different nuclei. Such a mycelium is said to be a **heterokaryon** (meaning "different nuclei"). In some species, the haploid nuclei pair off two to a cell, one from each parent. Such a mycelium is **dikaryotic** (meaning "two nuclei"). As a dikaryotic mycelium grows, the two nuclei in each cell divide in tandem without fusing. Because these cells retain two separate haploid nuclei, they differ from diploid cells, which have pairs of homologous chromosomes within a single nucleus.

Hours, days, or (in some fungi) even centuries may pass between plasmogamy and the next stage in the sexual cycle, **karyogamy**. During karyogamy, the haploid nuclei contributed by the two parents fuse, producing diploid cells. Zygotes and other transient structures form during karyogamy, the only diploid stage in most fungi. Meiosis then restores the haploid condition, ultimately leading to the formation of genetically diverse spores. Meiosis is a key step in sexual reproduction, so spores produced in this way are sometimes referred to as "sexual spores."

The sexual processes of karyogamy and meiosis generate extensive genetic variation, a prerequisite for natural selection. (See Concepts 13.2 and 23.1 to review how sex can increase genetic diversity.) The heterokaryotic condition also offers some of the advantages of diploidy in that one haploid genome may compensate for harmful mutations in the other.

Asexual Reproduction

Although many fungi can reproduce both sexually and asexually, some 20,000 species are only known to reproduce asexually. As with sexual reproduction, the processes of asexual reproduction vary widely among fungi.

Many fungi reproduce asexually by growing as filamentous fungi that produce (haploid) spores by mitosis; such species are informally referred to as **molds** if they form visible mycelia. Depending on your housekeeping habits, you may have observed molds in your kitchen, forming furry carpets on bread or fruit **(Figure 31.6)**. Molds typically grow

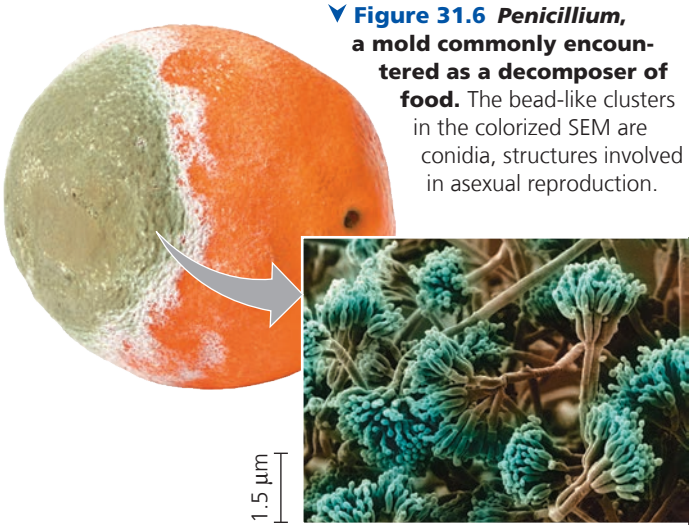
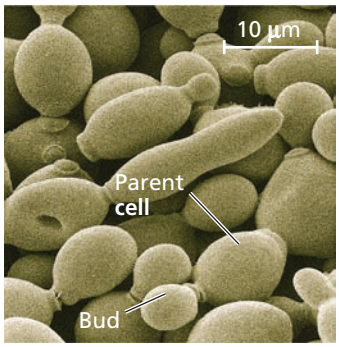


Figure 31.6 *Penicillium*, a mold commonly encountered as a decomposer of food. The bead-like clusters in the colorized SEM are conidia, structures involved in asexual reproduction.

rapidly and produce many spores asexually, enabling the fungi to colonize new sources of food. Many species that produce such spores can also reproduce sexually if they happen to contact a member of their species of a different mating type.

Other fungi reproduce asexually by growing as single-celled yeasts. Instead of producing spores, asexual reproduction in yeasts occurs by ordinary cell division or by the pinching of small "bud cells" off a parent cell **(Figure 31.7)**. As already mentioned, some fungi that grow as yeasts can also grow as filamentous mycelia.

Figure 31.7 The yeast *Saccharomyces cerevisiae* in several stages of budding (SEM).



Many yeasts and filamentous fungi have no known sexual stage in their life cycle. Since early mycologists (biologists who study fungi) classified fungi based mainly on their type of sexual structure, this posed a problem. Mycologists have traditionally lumped all fungi lacking sexual reproduction into a group called **deuteromycetes** (from the Greek *deutero*, second, and *mycete*, fungus). Whenever a sexual stage is discovered for a so-called deuteromycete, the species is reclassified in a particular phylum, depending on the type of sexual structures it forms. In addition to searching for sexual stages of such unassigned fungi, mycologists can now use genomic techniques to classify them.

CONCEPT CHECK 31.2

1. **MAKE CONNECTIONS ➤** Compare Figure 31.5 with Figure 13.6. In terms of haploidy versus diploidy, how do the life cycles of fungi and humans differ?
2. **WHAT IF? ➤** Suppose that you sample the DNA of two mushrooms on opposite sides of your yard and find that they are identical. Propose two hypotheses that could reasonably account for this result.

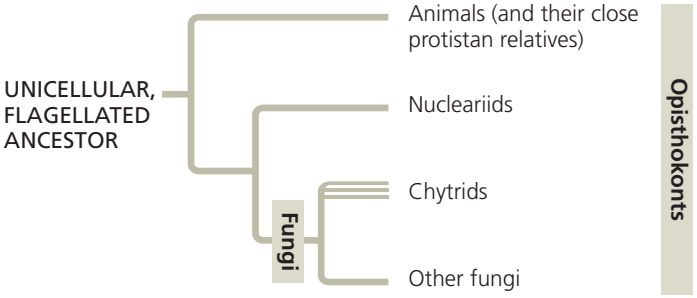
For suggested answers, see Appendix A.

CONCEPT 31.3

The ancestor of fungi was an aquatic, single-celled, flagellated protist

Data from both paleontology and molecular systematics offer insights into the early evolution of fungi. As a result, systematists now recognize that fungi and animals are more closely related to each other than either group is to plants or to most other eukaryotes.

Figure 31.8 Fungi and their close relatives. Molecular evidence indicates that the nucleariids, a group of single-celled protists, are the closest living relatives of fungi. The three parallel lines leading to the chytrids indicate that this group is paraphyletic.



The Origin of Fungi

Phylogenetic analyses suggest that fungi evolved from a flagellated ancestor. While the majority of fungi lack flagella, some of the earliest-diverging lineages of fungi (the chytrids, as we'll discuss shortly) do have flagella. Moreover, most of the protists that share a close common ancestor with animals and fungi also have flagella. DNA sequence data indicate that these three groups of eukaryotes—the fungi, the animals, and their protistan relatives—form a monophyletic group, or clade **(Figure 31.8)**. As discussed in Concept 28.5, members of this clade are called **opisthokonts**, a name that refers to the posterior (*opistho-*) location of the flagellum in these organisms.

DNA sequence data also indicate that fungi are more closely related to several groups of single-celled protists than they are to animals, suggesting that the ancestor of fungi was unicellular. One such group of unicellular protists, the **nucleariids**, consists of amoebas that feed on algae and bacteria. DNA evidence further indicates that animals are more closely related to a *different* group of protists (the choanoflagellates) than they are to either fungi or nucleariids. Together, these results suggest that multicellularity evolved in animals and fungi independently, from different single-celled ancestors.

Using molecular clock analyses, scientists have estimated that the ancestors of animals and fungi diverged into separate lineages more than a billion years ago. Fossils of certain unicellular, marine eukaryotes that lived as early as 1.5 billion years ago have been interpreted as fungi, but those claims remain controversial. Furthermore, although most scientists think that fungi originated in aquatic environments, the oldest fossils that are widely accepted as fungi are of terrestrial species that lived about 460 million years ago **(Figure 31.9)**. Overall, more

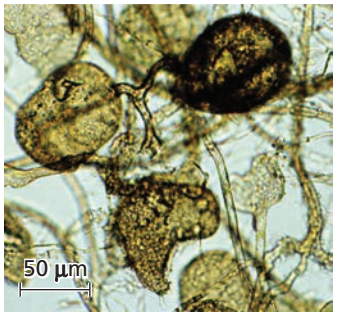


Figure 31.9 Fossil fungal hyphae and spores from the Ordovician period (about 460 million years ago) (LM).

fossils are needed to help clarify when fungi originated and what features were present in their earliest lineages.

Basal Fungal Groups

Insights into the nature of basal fungal groups have begun to emerge from recent genomic studies. For example, several studies have identified chytrids in the genus *Rozella* as having diverged from other fungi early in the history of the group. Furthermore, one metagenomics study placed *Rozella* within a large, previously unknown clade of unicellular fungi, tentatively called "cryptomycota." Like *Rozella* (and chytrids in general), fungi in the cryptomycota clade have flagellated spores. Current evidence also indicates that *Rozella* and other members of the cryptomycota are unique among fungi in that they do not synthesize a chitin-rich cell wall during any of their life cycle stages. This suggests that a cell wall strengthened by chitin—a key structural feature found in most fungi—may have arisen after the cryptomycota diverged from other fungi.

The Move to Land

Plants colonized land about 470 million years ago (see Concept 29.1), and fungi may well have colonized land before plants. Indeed, some researchers have described life on land before the arrival of plants as a "green slime" that consisted of cyanobacteria, algae, and a variety of small, heterotrophic species, including fungi. With their capacity for extracellular digestion, fungi would have been well suited for feeding on other early terrestrial organisms (or their remains).

Once on land, some fungi formed symbiotic associations with early plants. For example, 405-million-year-old fossils of the early plant *Aglaophyton* contain evidence of mycorrhizal relationships between plants and fungi (see Figure 25.12). This evidence includes fossils of hyphae that have penetrated within plant cells and formed structures that resemble the arbuscules formed today by arbuscular mycorrhizae. Similar structures have been found in a variety of other early plants, suggesting that plants probably existed in beneficial relationships with fungi from the earliest periods of colonization of land. The earliest plants lacked roots, limiting their ability to extract nutrients from the soil. As occurs in mycorrhizal associations today, it is likely that soil nutrients were transferred to early plants via the extensive mycelia formed by their symbiotic fungal partners.

Support for the antiquity of mycorrhizal associations has also come from recent molecular studies. For a mycorrhizal fungus and its plant partner to establish a symbiotic relationship, certain genes must be expressed by the fungus and other genes must be expressed by the plant. Researchers focused on three plant genes (called *sym* genes) whose expression is required for the formation of mycorrhizae in flowering plants. They found that these genes were present in all major plant lineages, including basal lineages such as liverworts (see Figure 29.8). Furthermore, after they transferred a liverwort *sym* gene to a flowering plant mutant that could not form mycorrhizae, the mutant recovered its ability to form mycorrhizae. These results suggest that mycorrhizal *sym* genes were present in early plants—and that the function of these genes has been conserved for hundreds of millions of years as plants continued to adapt to life on land.

CONCEPT CHECK 31.3

1. Why are fungi classified as opisthokonts despite the fact that most fungi lack flagella?
2. Describe the importance of mycorrhizae, both today and in the colonization of land. What evidence supports the antiquity of mycorrhizal associations?
3. **WHAT IF?** ➤ If fungi colonized land before plants, where might the fungi have lived? How would their food sources have differed from what they feed on today?

For suggested answers, see Appendix A.

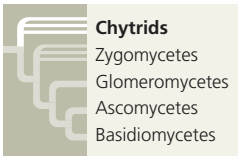
CONCEPT 31.4

Fungi have radiated into a diverse set of lineages

In the past decade, molecular analyses have helped clarify the evolutionary relationships between fungal groups, although there are still areas of uncertainty. **Figure 31.10** presents a simplified version of one current hypothesis. In this section, we will survey each of the major fungal groups identified in this phylogenetic tree.

The fungal groups shown in Figure 31.10 may represent only a small fraction of the diversity of extant fungal groups. (Extant lineages are those that have surviving members.) While there are roughly 100,000 known species of fungi, scientists have estimated that the actual diversity may be closer to 1.5 million species. Two recent metagenomic studies support such higher estimates: the cryptomycota (see Concept 31.3) and other entirely new groups of unicellular fungi were discovered, and the genetic variation found in some of these groups is as large as that found across all of the groups shown in Figure 31.10.

Chytrids



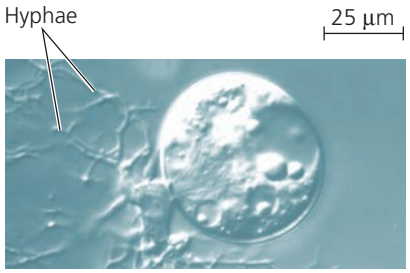
The fungi classified in the phylum Chytridiomycota, called **chytrids**, are ubiquitous in lakes and soil, and as described in several recent metagenomic studies, more than 20 new clades of chytrids have been found in hydrothermal vent and other marine communities. Some of the approximately 1,000 chytrid species are decomposers, while others are parasites of protists, other fungi, plants, or animals; as we'll see later in the chapter, one such chytrid parasite has likely contributed

Figure 31.10 Exploring Fungal Diversity

Many mycologists currently recognize five major groups of fungi, although recent genomic evidence indicates that the chytrids and zygomycetes are paraphyletic (as indicated by the parallel lines).

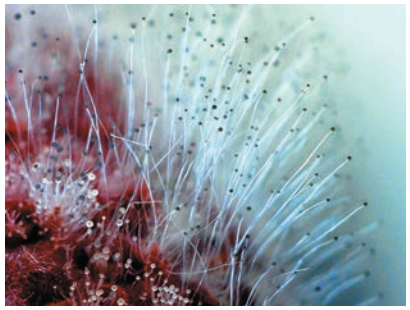
Chytrids (1,000 species)

In chytrids such as *Chytriomyces*, the globular fruiting body forms multicellular, branched hyphae (LM); other species are single-celled. Ubiquitous in lakes and soil, chytrids have flagellated spores and are thought to include some of the earliest fungal groups to diverge from other fungi.



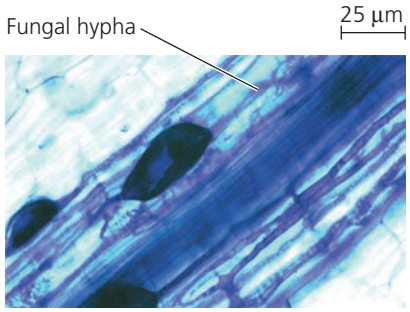
Zygomycetes (1,000 species)

The hyphae of some zygomycetes, including this mold in the genus *Mucor* (LM), grow rapidly on foods such as fruits and bread. As such, the fungi may act as decomposers (if the food is not alive) or parasites; other species live as neutral (commensal) symbionts.



Glomeromycetes (160 species)

The glomeromycetes form arbuscular mycorrhizae with plant roots, supplying minerals and other nutrients to the roots; more than 80% of all plant species have such mutualistic partnerships with glomeromycetes. This LM shows glomeromycete hyphae (filaments stained dark blue) within a plant root.



Ascomycetes (65,000 species)

Also called sac fungi, members of this diverse group are common to many marine, freshwater, and terrestrial habitats. The cup-shaped ascocarp (fruiting body) of the ascomycete shown here (*Aleuria aurantia*) gives this species its common name: orange peel fungus.

Basidiomycetes (30,000 species)

Often important as decomposers and ectomycorrhizal fungi, basidiomycetes, or club fungi, are unusual in having a long-lived, heterokaryotic stage in which each cell has two nuclei (one from each parent). The fruiting bodies—commonly called mushrooms—of this fly agaric (*Amanita muscaria*) are a familiar sight in coniferous forests of the Northern Hemisphere.

to the global decline of amphibian populations. Still other chytrids are important mutualists. For example, anaerobic chytrids that live in the digestive tracts of sheep and cattle help to break down plant matter, thereby contributing significantly to the animal's growth.

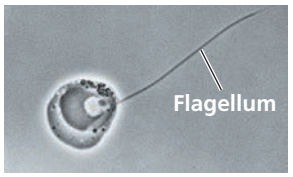


Figure 31.11 Flagellated chytrid zoospore (TEM).

As discussed earlier, molecular evidence indicates that some chytrid lineages diverged early in fungal evolution. The fact that chytrids are unique among fungi in having flagellated spores, called **zoospores** **(Figure 31.11)**, supports this hypothesis. Like other fungi, chytrids (other than those in the recently discovered cryptomycota clade) have cell walls made of chitin, and they also share certain key enzymes and metabolic pathways with other fungal groups. Some chytrids form colonies with hyphae, while others exist as single spherical cells.

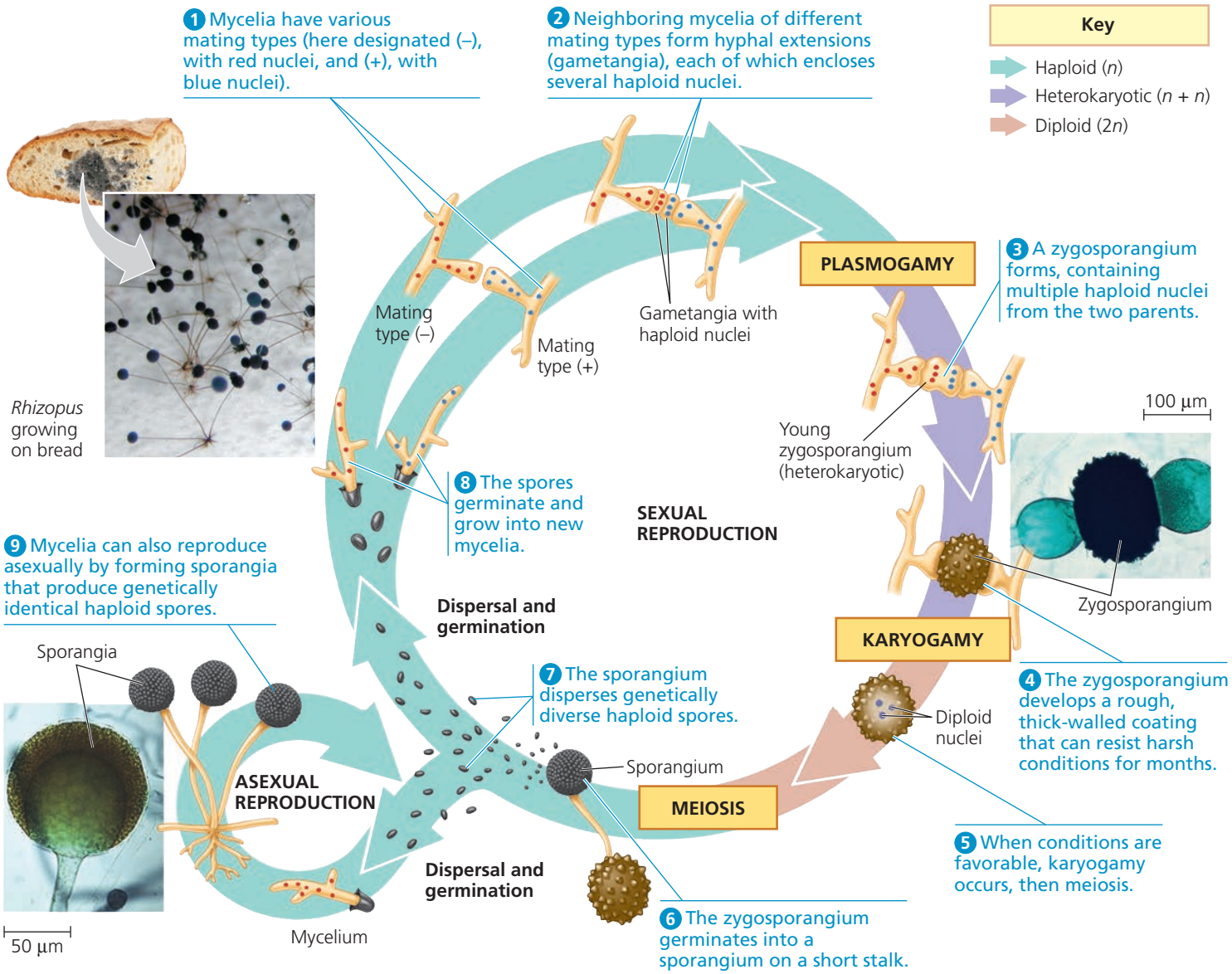
Video: *Phlyctochytrium* Zoospore Release

Zygomycetes

There are approximately 1,000 known species of **zygomycetes**, fungi in the phylum Zygomycota. This diverse phylum includes species of fast-growing molds responsible for causing foods such as bread, peaches, strawberries, and sweet potatoes to rot during storage. Other zygomycetes live as parasites or as commensal (neutral) symbionts of animals.

The life cycle of *Rhizopus stolonifer* (black bread mold) is fairly typical of zygomycete species **(Figure 31.12)**. Its hyphae

Figure 31.12 The life cycle of the zygomycete *Rhizopus stolonifer* (black bread mold).



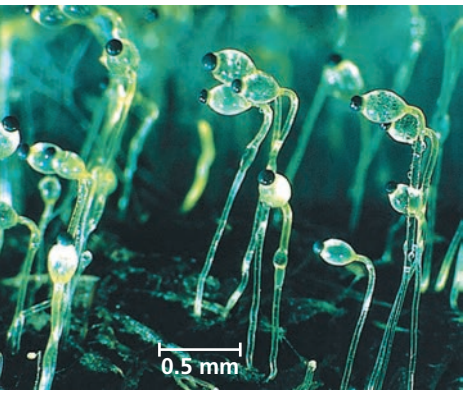


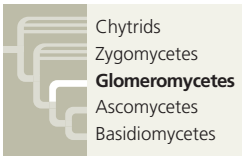
Figure 31.13 *Pilobolus* aiming its sporangia. This zygomycete decomposes animal dung. Its spore-bearing hyphae bend toward light, where there are likely to be openings in the vegetation through which spores may reach fresh grass. The fungus then launches its sporangia in a jet of water that can travel up to 2.5 m. Grazing animals ingest the fungi with the grass and then scatter the spores in feces, thereby enabling the next generation of fungi to grow.

spread out over the food surface, penetrate it, and absorb nutrients. The hyphae are coenocytic, with septa found only where reproductive cells are formed. In the asexual phase, bulbous black sporangia develop at the tips of upright hyphae. Within each sporangium, hundreds of genetically identical haploid spores develop and are dispersed through the air. Spores that happen to land on moist food germinate, growing into new mycelia.

If environmental conditions deteriorate—for instance, if the mold consumes all its food—*Rhizopus* may reproduce sexually. The parents in a sexual union are mycelia of different mating types, which possess different chemical markers but may appear identical. Plasmogamy produces a sturdy structure called a **zygosporangium** (plural, *zygosporangia*), in which karyogamy and then meiosis occur. Note that while a zygosporangium represents the zygote ($2n$) stage in the life cycle, it is not a zygote in the usual sense (that is, a cell with one diploid nucleus). Rather, a zygosporangium is a multinucleate structure, first heterokaryotic with many haploid nuclei from the two parents, then with many diploid nuclei after karyogamy.

Zygosporangia are resistant to freezing and drying and are metabolically inactive. When conditions improve, the nuclei of the zygosporangium undergo meiosis, the zygosporangium germinates into a sporangium, and the sporangium releases genetically diverse haploid spores that may colonize a new substrate. Some zygomycetes, such as *Pilobolus*, can actually "aim" and then shoot their sporangia toward bright light **(Figure 31.13)**.

Glomeromycetes



The **glomeromycetes**, fungi assigned to the phylum Glomeromycota, were formerly thought to be zygomycetes. But recent molecular studies, including a phylogenetic analysis of DNA sequence data from hundreds of fungal species, indicate that glomeromycetes form a separate clade. Although only 200 species have been described to date, molecular studies indicate that the actual number of species may be much

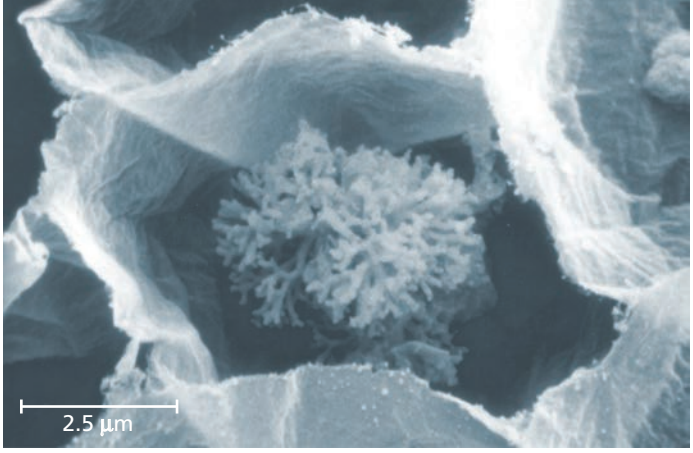
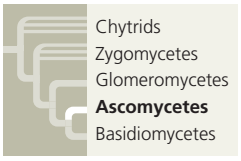


Figure 31.14 Arbuscular mycorrhizae. Most glomeromycetes form arbuscular mycorrhizae with plant roots, supplying minerals and other nutrients to the roots. This SEM depicts the branched hyphae—an arbuscule—of *Glomus mosseae* bulging into a root cell by pushing in the membrane (the root has been treated to remove the cytoplasm).

higher. The glomeromycetes are an ecologically significant group in that nearly all of them form arbuscular mycorrhizae **(Figure 31.14)**. The tips of the hyphae that push into plant root cells branch into tiny treelike arbuscules. More than 80% of all plant species have such mutualistic partnerships with glomeromycetes.

Ascomycetes



Mycologists have described 65,000 species of **ascomycetes**, fungi in the phylum Ascomycota, from a wide variety of marine, freshwater, and terrestrial habitats. The defining feature of ascomycetes is the production of spores (called ascospores) in saclike **asci** (singular, *ascus*); thus, they are commonly called *sac fungi*. During their sexual stage, most ascomycetes develop fruiting bodies, called **ascocarps**, which range in size from microscopic to macroscopic **(Figure 31.15)**. The ascocarps contain the spore-forming asci.

Figure 31.15 Ascomycetes (sac fungi).

▶ *Tuber melanosporum* is a truffle species that forms ectomycorrhizae with trees. The ascocarp grows underground and emits a strong odor. These ascocarps have been dug up and the middle one sliced open.

◀ The edible ascocarp of *Morchella esculenta*, the tasty morel, is often found under trees in orchards.

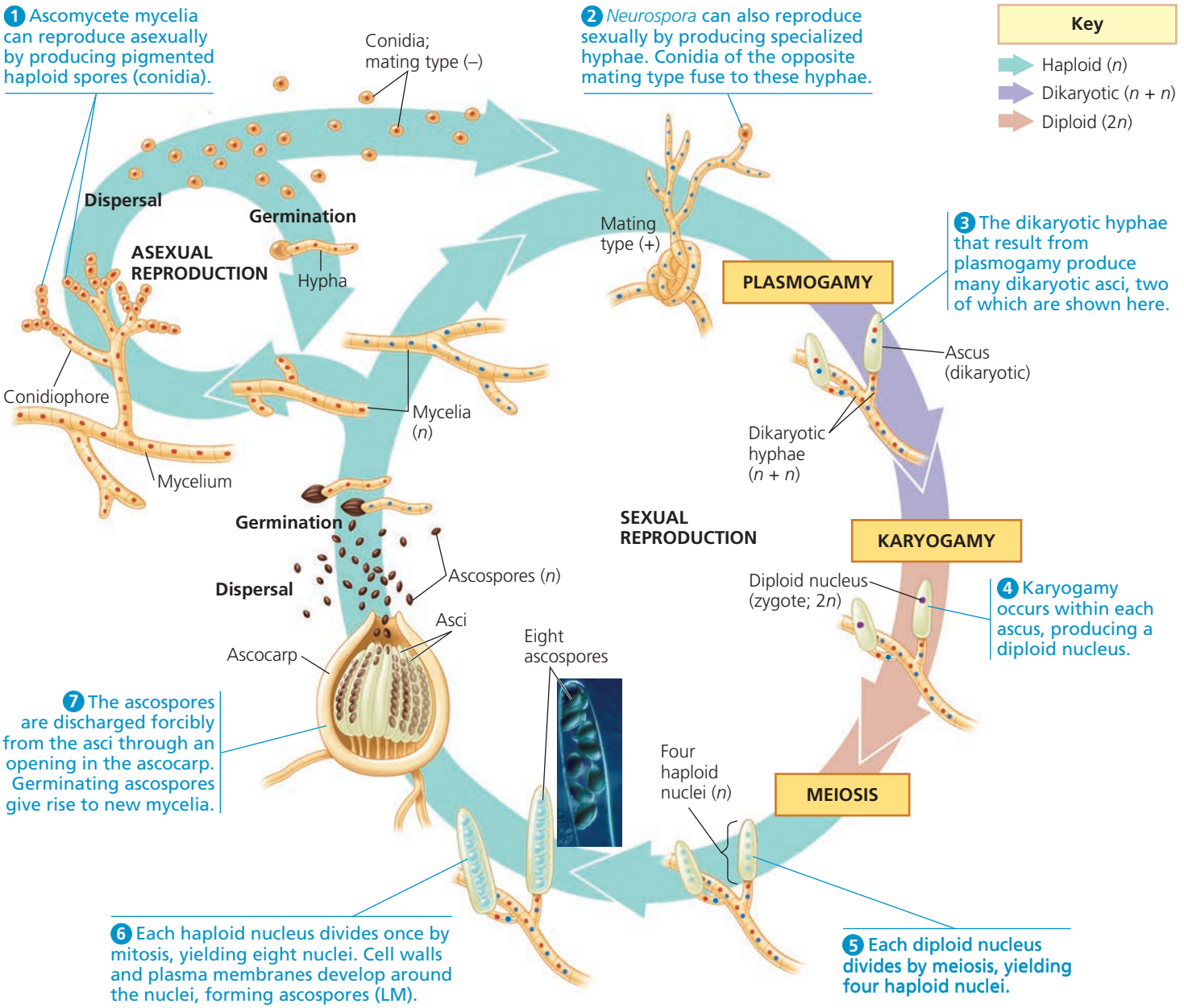
? *Ascomycetes vary greatly in morphology (see also Figure 31.10). How could you confirm that a fungus is an ascomycete?*

Ascomycetes vary in size and complexity from unicellular yeasts to elaborate cup fungi and morels (see Figure 31.15). They include some of the most devastating plant pathogens, which we will discuss later. However, many ascomycetes are important decomposers, particularly of plant material. More than 25% of all ascomycete species live with green algae or cyanobacteria in beneficial symbiotic associations called lichens. Some ascomycetes form mycorrhizae with plants. Many others live between mesophyll cells in leaves; some of these species release toxic compounds that help protect the plant from insects.

Although the life cycles of various ascomycete groups differ in the details of their reproductive structures and processes, we'll illustrate some common elements using the bread mold *Neurospora crassa* **(Figure 31.16)**. Ascomycetes reproduce asexually by producing enormous numbers of asexual spores called **conidia** (singular, *conidium*). Unlike the asexual spores of most zygomycetes, in most ascomycetes, conidia are not formed inside sporangia. Rather, they are produced externally at the tips of specialized hyphae called conidiophores, often in clusters or long chains, from which they may be dispersed by the wind.

Conidia may also be involved in sexual reproduction, fusing with hyphae from a mycelium of a different mating type, as occurs in *Neurospora*. Fusion of two different mating types is followed by plasmogamy, resulting in the formation

Figure 31.16 The life cycle of *Neurospora crassa*, an ascomycete. *Neurospora* is a bread mold and research organism that also grows in the wild on burned vegetation.



VISUAL SKILLS ➤ *What is the ploidy of a cell in the specialized hypha shown in 2?*

of dikaryotic cells, each with two haploid nuclei representing the two parents. The cells at the tips of these dikaryotic hyphae develop into many asci. Within each ascus, karyogamy combines the two parental genomes, and then meiosis forms four genetically different nuclei. This is usually followed by a mitotic division, forming eight ascospores. The ascospores develop in and are eventually discharged from the ascocarp.

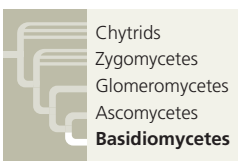
Compared to the life cycle of zygomycetes, the extended dikaryotic stage of ascomycetes (and also basidiomycetes) provides additional opportunities for genetic recombination. In *Neurospora*, for example, many dikaryotic cells can develop into asci. The haploid nuclei in these asci fuse, and their genomes then recombine during meiosis, resulting in a multitude of genetically different offspring from one mating event (see steps 3–5 in Figure 31.16).

As described in Figure 17.2, biologists in the 1930s used *Neurospora* in research that led to the one gene–one enzyme hypothesis. Today, this ascomycete continues to serve as a model research organism. In 2003, its entire genome was published. This tiny fungus has about three-fourths as many genes as the fruit fly *Drosophila* and about half as many as a human **(Table 31.1)**. The *Neurospora* genome is relatively compact, having few of the stretches of noncoding DNA that occupy so much space in the genomes of humans and many other eukaryotes. In fact, there is evidence that *Neurospora* has a genomic defense system that prevents noncoding DNA such as transposons from accumulating.

Table 31.1 Comparison of Gene Density in *Neurospora*, *Drosophila*, and *Homo sapiens*

	Genome Size (million base pairs)	Number of Genes	Gene Density (genes per million base pairs)
Neurospora crassa (ascomycete fungus)	41	9,700	236
Drosophila melanogaster (fruit fly)	165	14,000	85
Homo sapiens (human)	3,000	<21,000	7

Basidiomycetes



About 30,000 species, including mushrooms, puffballs, and shelf fungi, are called **basidiomycetes** and are classified in the phylum Basidiomycota **(Figure 31.17)**. This phylum also includes mutualists that form mycorrhizae and two groups of destructive plant parasites: rusts and smuts. The name of the phylum derives from the **basidium** (plural, *basidia*; Latin for "little pedestals"), a cell in which karyogamy occurs, followed immediately by meiosis. The club-like shape of the basidium also gives rise to the common name *club fungus*.

Basidiomycetes are important decomposers of wood and other plant material. Of all the fungi, certain basidiomycetes are the best at decomposing the complex polymer lignin, an abundant component of wood. Many shelf fungi break down the wood of weak or damaged trees and continue to decompose the wood after the tree dies.

The life cycle of a basidiomycete usually includes a long-lived dikaryotic mycelium. As in ascomycetes, this extended

▶ Shelf fungi, important decomposers of wood

◀ Puffballs emitting spores

▶ Maiden veil fungus (*Dictyphora*), a fungus with an odor like rotting meat

Figure 31.17 Basidiomycetes (club fungi).

dikaryotic stage provides many opportunities for genetic recombination events, in effect multiplying the result of a single mating. Periodically, in response to environmental stimuli, the mycelium reproduces sexually by producing elaborate fruiting bodies called **basidiocarps** **(Figure 31.18)**. The common white mushrooms in the supermarket are familiar examples of a basidiocarp.

By concentrating growth in the hyphae of mushrooms, a basidiomycete mycelium can erect its fruiting structures in just a few hours; a mushroom pops up as it absorbs water and as cytoplasm streams in from the dikaryotic mycelium. By this process, in some species a ring of mushrooms, popularly called a "fairy ring," may appear literally overnight **(Figure 31.19)**. The mycelium below the fairy ring expands outward at a rate of about 30 cm per year, decomposing organic matter in the soil as it grows. Some giant fairy rings are produced by mycelia that are centuries old.

After a mushroom forms, its cap supports and protects a large surface area of dikaryotic basidia on gills. During karyogamy, the two nuclei in each basidium fuse, producing a diploid nucleus (see Figure 31.18). This nucleus then undergoes meiosis, yielding four haploid nuclei, each of which ultimately develops into a basidiospore. Large numbers of basidiospores are produced: The gills of a common white mushroom have a surface area of about 200 cm^2 and may drop a billion basidiospores, which blow away.

Figure 31.18 The life cycle of a mushroom-forming basidiomycete.

VISUAL SKILLS ➤ *Use the diagram to determine the ploidy of a cell in the aboveground stalk of a mushroom.*

Animation: Life Cycle of a Mushroom

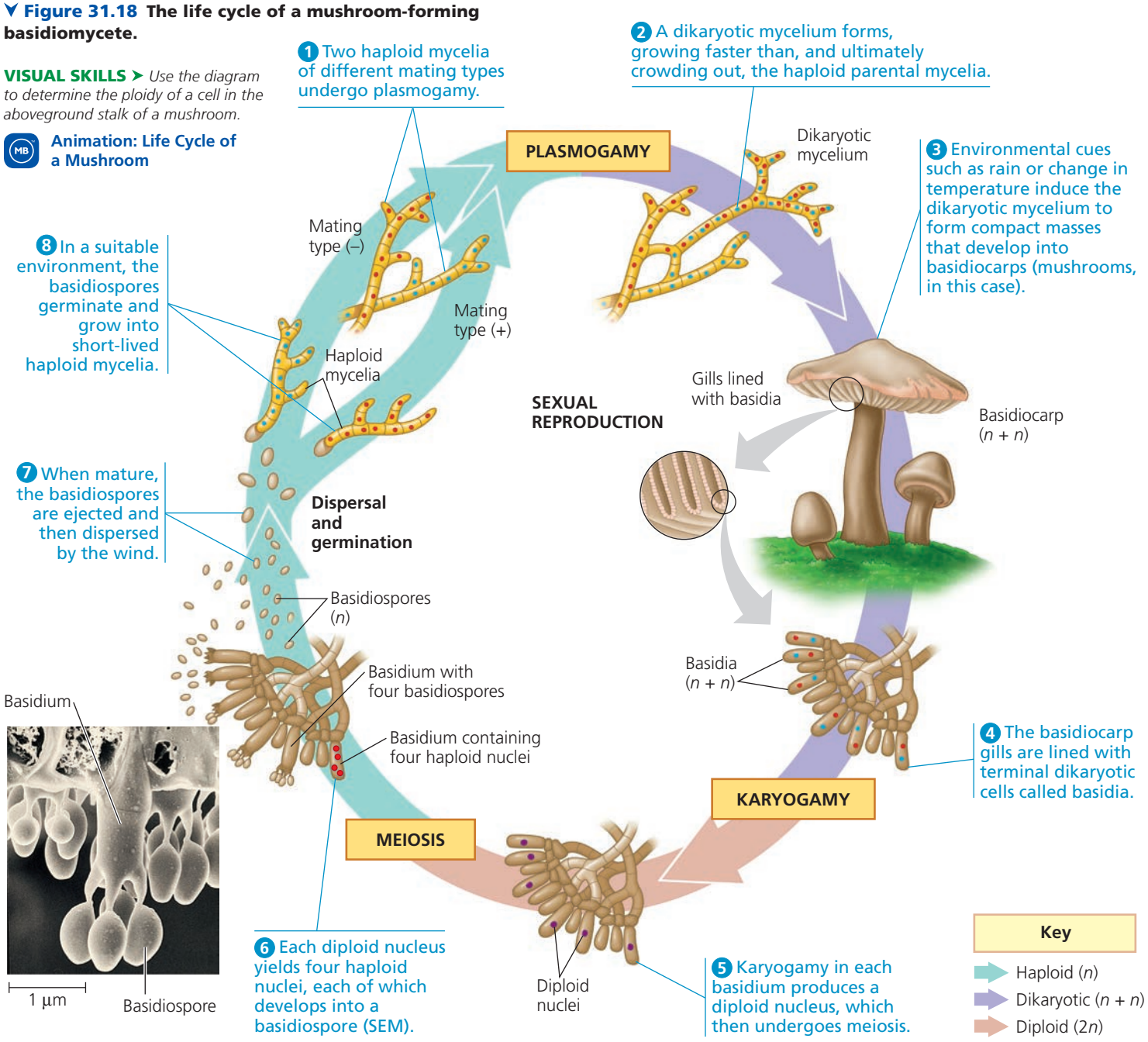


Figure 31.19 A fairy ring. According to legend, mushroom rings spring up where fairies have danced on a moonlit night. The text provides a biological explanation of how these rings form.

CONCEPT CHECK 31.4

1. What feature of chytrids supports the hypothesis that they include members of basal fungal lineages?
2. Give examples of how form fits function in zygomycetes, glomeromycetes, ascomycetes, and basidiomycetes.
3. **WHAT IF?** ➤ Suppose that the mutation of an ascomycete changed its life cycle so that plasmogamy, karyogamy, and meiosis occurred in quick succession. How might this affect the ascospores and ascocarps?

For suggested answers, see Appendix A.

CONCEPT 31.5

Fungi play key roles in nutrient cycling, ecological interactions, and human welfare

In our survey of fungal classification, we've touched on some of the ways fungi influence other organisms. We will now look more closely at these impacts, focusing on how fungi act as decomposers, mutualists, and pathogens.

Fungi as Decomposers

Fungi are well adapted as decomposers of organic material, including the cellulose and lignin of plant cell walls. In fact, almost any carbon-containing substrate—even jet fuel and house paint—can be consumed by at least some fungi. The same is true of bacteria. As a result, fungi and bacteria are primarily responsible for keeping ecosystems stocked with the inorganic nutrients essential for plant growth. Without these decomposers, carbon, nitrogen, and other elements would remain tied up in organic matter. If that were to happen, plants and the animals that eat them could not exist because elements taken from the soil would not be returned. Without decomposers, life as we know it would cease.

Fungi as Mutualists

Fungi may form mutualistic relationships with plants, algae, cyanobacteria, and animals. Mutualistic fungi absorb nutrients from a host organism, but they reciprocate with actions that benefit the host—as we already saw for the key mycorrhizal associations that fungi form with most vascular plants.

Fungus-Plant Mutualisms

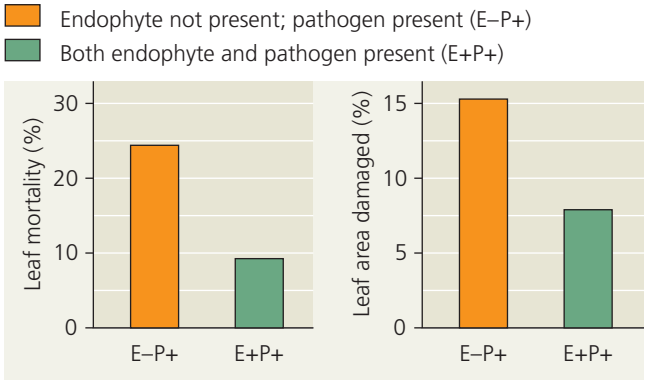
Along with mycorrhizal fungi, all plant species studied to date appear to harbor symbiotic **endophytes**, fungi (or bacteria) that live inside leaves or other plant parts without causing harm. Most fungal endophytes identified to date are ascomycetes. Fungal endophytes benefit certain grasses and other nonwoody plants by making toxins that deter herbivores or by increasing host plant tolerance of heat, drought, or heavy metals. As described in **Figure 31.20**, researchers

Figure 31.20

Inquiry Do fungal endophytes benefit a woody plant?

Experiment Fungal endophytes are symbiotic fungi found within the bodies of all plants examined to date. A. Elizabeth Arnold, at the University of Arizona, Tucson, and colleagues tested whether fungal endophytes benefit the cacao tree (*Theobroma cacao*). This tree, whose name means "food of the gods" in Greek, is the source of the beans used to make chocolate, and it is cultivated throughout the tropics. A particular mixture of fungal endophytes was added to the leaves of some cacao seedlings, but not others. (In cacao, fungal endophytes colonize leaves after the seedling germinates.) The seedlings were then inoculated with a virulent pathogen, the protist *Phytophthora*.

Results Fewer leaves were killed by the pathogen in seedlings with fungal endophytes than in seedlings without endophytes. Among leaves that survived, pathogens damaged less of the leaf surface area in seedlings with endophytes than in seedlings without endophytes.



Conclusion The presence of endophytes appears to benefit cacao trees by reducing the leaf mortality and damage caused by *Phytophthora*.

Data from A. E. Arnold et al., Fungal endophytes limit pathogen damage in a tropical tree, *Proceedings of the National Academy of Sciences* 100:15649–15654 (2003).

WHAT IF? ➤ *Arnold and colleagues also performed control treatments. Suggest two controls they might have used, and explain how each would be helpful in interpreting the results described here.*

Figure 31.21 Fungus-gardening insects. These leaf-cutting ants depend on fungi to convert plant material to a form the insects can digest. The fungi, in turn, depend on the nutrients from the leaves the ants feed them.

studying how fungal endophytes affect a woody plant tested whether leaf endophytes benefit seedlings of the cacao tree, *Theobroma cacao*. Their findings show that the fungal endophytes of woody flowering plants can play an important role in defending against pathogens.

Fungus-Animal Mutualisms

As mentioned earlier, some fungi share their digestive services with animals, helping break down plant material in the guts of cattle and other grazing mammals. Many species of ants take advantage of the digestive power of fungi by raising them in "farms." Leaf-cutter ants, for example, scour tropical forests in search of leaves, which they cannot digest on their own but carry back to their nests and feed to the fungi **(Figure 31.21)**. As the fungi grow, their hyphae develop specialized swollen tips that are rich in proteins and carbohydrates. The ants feed primarily on these nutrient-rich tips. Not only do the fungi break down plant leaves into substances the insects can digest, but they also detoxify plant defensive compounds that would otherwise kill or harm the ants. In some tropical forests, the fungi have helped these insects become the major consumers of leaves.

The evolution of such farmer ants and that of their fungal "crops" have been tightly linked for over 50 million years. The fungi have become so dependent on their caretakers that in many cases they can no longer survive without the ants, and vice versa.

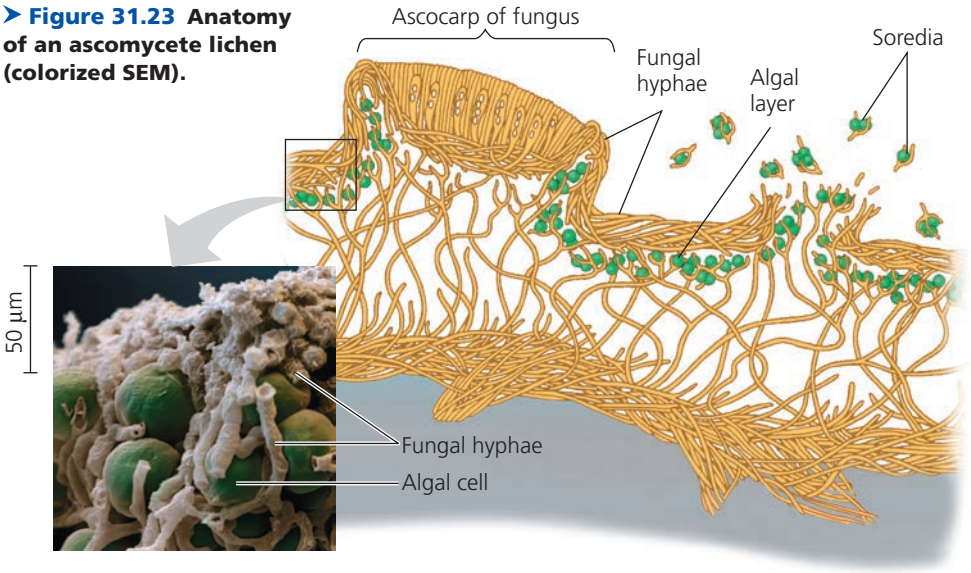
Lichens

A **lichen** is a symbiotic association between a photosynthetic microorganism and a fungus in which millions of photosynthetic cells are held in a mass of fungal hyphae. Lichens grow on the surfaces of rocks, rotting logs, trees, and roofs in various forms **(Figure 31.22)**. The photosynthetic partners are unicellular or filamentous green algae or cyanobacteria. The fungal component is most often an ascomycete, but one glomeromycete and 75 basidiomycete lichens are known. The fungus usually gives a lichen its overall shape and structure, and tissues formed by hyphae account for most of the lichen's mass. The cells of the alga or cyanobacterium generally occupy an inner layer below the lichen surface **(Figure 31.23)**.

Figure 31.22 Variation in lichen growth forms.

The merger of fungus and alga or cyanobacterium is so complete that lichens are given scientific names as though they were single organisms; to date, 17,000 lichen species have been described. As might be expected of such "dual organisms," asexual reproduction as a symbiotic unit is common. This can occur either by fragmentation of the parental lichen or by the formation of **soredia** (singular, *soredium*), small clusters of hyphae with embedded algae (see Figure 31.23). The fungi of many lichens also reproduce sexually.

In most lichens, each partner provides something the other could not obtain on its own. The alga or cyanobacterium provides carbon compounds; a cyanobacterium also



➤ Figure 31.23 Anatomy of an ascomycete lichen (colorized SEM).

fixes nitrogen (see Concept 27.3) and provides organic nitrogen compounds. The fungus provides its photosynthetic partner with a suitable environment for growth. The physical arrangement of hyphae allows for gas exchange, protects the photosynthetic partner, and retains water and minerals, most of which are absorbed from airborne dust or from rain. The fungus also secretes acids, which aid in the uptake of minerals.

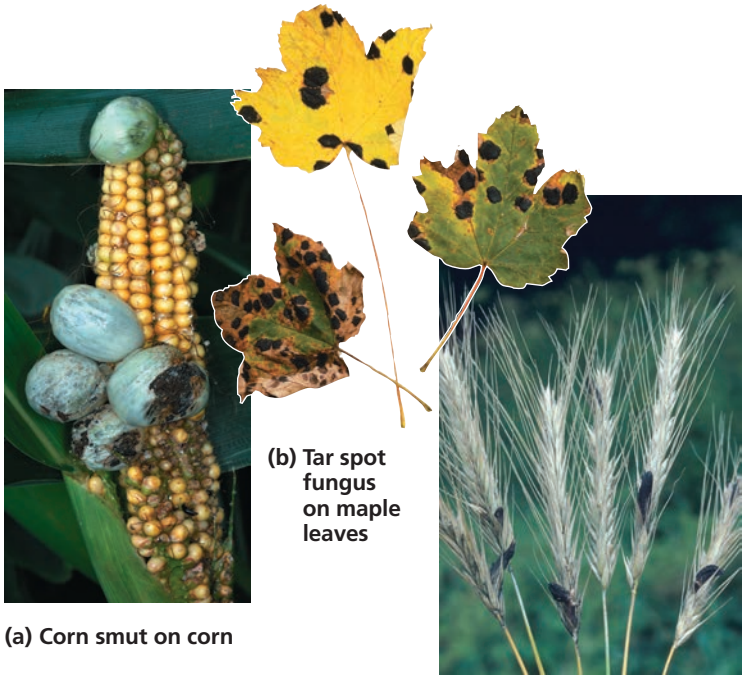
Lichens are important pioneers on cleared rock and soil surfaces, such as volcanic flows and burned forests. They break down the surface by physically penetrating and chemically attacking it, and they trap windblown soil. Nitrogen-fixing lichens also add organic nitrogen to some ecosystems. These processes make it possible for a succession of plants to grow. Fossils show that lichens were on land 420 million years ago. These early lichens may have modified rocks and soil much as they do today, helping pave the way for plants.

Fungi as Parasites

Like mutualistic fungi, parasitic fungi absorb nutrients from the cells of living hosts, but they provide no benefits in return. About 30% of the 100,000 known species of fungi make a living as parasites or pathogens, mostly of plants **(Figure 31.24)**. An example of a plant pathogen is *Cryphonectria parasitica*, the ascomycete fungus that causes chestnut blight, which dramatically changed the landscape of the northeastern United States. Accidentally introduced via trees imported from Asia in the early 1900s, spores of the fungus entered cracks in the bark of American chestnut trees and produced hyphae, killing many trees. The once-common chestnuts now survive mainly as sprouts from the stumps of former trees. Another ascomycete, *Fusarium circinatum*, causes pine pitch canker, a disease that threatens pines throughout the world. Between 10% and 50% of the world's fruit harvest is lost annually due to fungi, and grain crops also suffer major losses each year.

Some fungi that attack food crops produce compounds that are toxic to humans. One example is the ascomycete *Claviceps purpurea*, which grows on rye plants, forming purple structures called ergots (see Figure 31.24c). If infected rye is milled into flour, toxins from the ergots can cause ergotism, characterized by gangrene, nervous spasms, burning sensations, hallucinations, and temporary insanity. An epidemic of ergotism around 944 CE killed up to 40,000 people in France. One compound that has been isolated from ergots is lysergic acid, the raw material from which the hallucinogen LSD is made.

Although animals are less susceptible to parasitic fungi than are plants, about 500 fungi are known to parasitize animals. One such parasite, the chytrid *Batrachochytrium dendrobatidis*, has been implicated in the recent decline or extinction of about 200 species of frogs and other amphibians



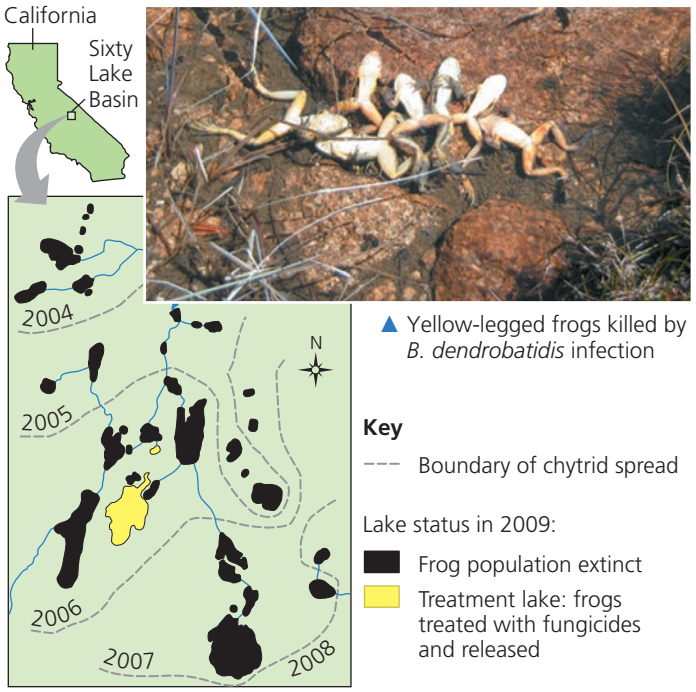
(a) Corn smut on corn

(b) Tar spot fungus on maple leaves

(c) Ergots on rye

➤ Figure 31.24 Examples of fungal diseases of plants.

Figure 31.25 Amphibians under attack. Could a fungal parasite have caused some of the many declines and extinctions of amphibian populations in recent decades? One study found that the number of yellow-legged frogs (*Rana muscosa*) plummeted after the chytrid *Batrachochytrium dendrobatidis* reached the Sixty Lake Basin area of California. In the years leading up to the chytrid's 2004 arrival, there had been more than 2,300 frogs in these lakes. By 2009, only 38 frogs remained; all the survivors were in two lakes (yellow) where frogs had been treated with a fungicide to reduce the chytrid's impact.



▲ Yellow-legged frogs killed by *B. dendrobatidis* infection

INTERPRET THE DATA ➤ *Do the data depicted indicate that the chytrid caused or is correlated to the drop in frog numbers? Explain.*

(Figure 31.25). This chytrid can cause severe skin infections, leading to massive die-offs. Field observations and studies of museum specimens indicate that *B. dendrobatidis* first appeared in frog populations shortly before their declines in Australia, Costa Rica, the United States, and other countries. In addition, in regions where it infects frogs, this chytrid has very low levels of genetic diversity. These findings are consistent with the hypothesis that *B. dendrobatidis* has emerged recently and spread rapidly across the globe, decimating many amphibian populations.

The general term for an infection in an animal by a fungal parasite is **mycosis**. In humans, skin mycoses include the disease ringworm, so named because it appears as circular red areas on the skin. Most commonly, the ascomycetes that cause ringworm grow on the feet, causing the intense itching and blisters known as athlete's foot. Though highly contagious, athlete's foot and other ringworm infections can be treated with fungicidal lotions and powders.

Systemic mycoses, by contrast, spread through the body and usually cause very serious illnesses. They are typically caused by inhaled spores. For example, coccidioidomycosis is a systemic mycosis that produces tuberculosis-like symptoms in the lungs. Each year, hundreds of cases in North America require treatment with antifungal drugs, without which the disease could be fatal.

Some mycoses are opportunistic, occurring only when a change in the body's microorganisms, chemical environment, or immune system allows fungi to grow unchecked. *Candida albicans*, for example, is one of the normal inhabitants of moist epithelia, such as the vaginal lining. Under certain circumstances, *Candida* can grow too rapidly and become pathogenic, leading to so-called "yeast infections." Many other opportunistic mycoses in humans have become more common in recent decades, due in part to AIDS, which compromises the immune system.

Practical Uses of Fungi

The dangers posed by fungi should not overshadow their immense benefits. We depend on their ecological services as decomposers and recyclers of organic matter. In addition, mushrooms are not the only fungi of interest for human consumption. Fungi are used to ripen Roquefort and other blue cheeses. Morels and truffles, the edible fruiting bodies of various ascomycetes, are highly prized for their complex flavors (see Figure 31.15). These fungi can sell for hundreds to thousands of dollars a pound. Truffles release strong odors that attract mammals and insects, which in nature feed on them and disperse their spores. In some cases, the odors mimic the pheromones (sex attractants) of certain mammals. For example, the odors of several European truffles mimic the pheromones released by male pigs, which explains why truffle hunters sometimes use female pigs to help find these delicacies.

Humans have used yeasts to produce alcoholic beverages and bread for thousands of years. Under anaerobic conditions, yeasts ferment sugars to alcohol and CO_2, which causes dough to rise. Only relatively recently have the yeasts involved been separated into pure cultures for more controlled use. The yeast *Saccharomyces cerevisiae* is the most important of all cultured fungi (see Figure 31.7). It is available as many strains of baker's yeast and brewer's yeast.

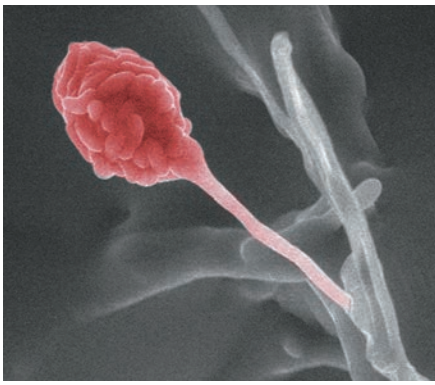
Many fungi have great medical value as well. For example, a compound extracted from ergots is used to reduce high blood pressure and to stop maternal bleeding after childbirth. Some fungi produce antibiotics that are effective in treating bacterial infections. In fact, the first antibiotic discovered was penicillin, made by the ascomycete mold *Penicillium*. Other examples of pharmaceuticals derived from fungi include cholesterol-lowering drugs and cyclosporine, a drug used to suppress the immune system after organ transplants.

Fungi also figure prominently in basic research. For example, the yeast *Saccharomyces cerevisiae* is used to study the

molecular genetics of eukaryotes because its cells are easy to culture and manipulate. Scientists are gaining insight into the genes involved in Parkinson's disease by examining the functions of homologous genes in *S. cerevisiae*.

Genetically modified fungi also hold much promise. For example, scientists have succeeded in engineering a strain of *S. cerevisiae* that produces human glycoproteins, including insulin-like growth factor. Such fungus-produced glycoproteins have the potential to treat people with medical conditions that prevent them from producing these compounds. Meanwhile, other researchers are sequencing the genome of *Gliocladium roseum*, an ascomycete that can grow on wood or agricultural waste and that naturally produces hydrocarbons similar to those in diesel fuel **(Figure 31.26)**. They hope to decipher the metabolic pathways by which *G. roseum* synthesizes hydrocarbons, with the goal of harnessing those pathways to produce biofuels without reducing land area for growing food crops (as occurs when ethanol is produced from corn).

Having now completed our survey of the kingdom Fungi, we will turn in the rest of this unit to the closely related kingdom Animalia, to which we humans belong.



➤ Figure 31.26 Can this fungus be used to produce biofuels? The ascomycete *Gliocladium roseum* can produce hydrocarbons similar to those in diesel fuel (colorized SEM).

CONCEPT CHECK 31.5

1. Describe the mutualistic relationship between leaf-cutter ants and fungi.
2. How can a benign fungus residing in the human body turn into a pathogenic one?
3. **WHAT IF? ➤** How might life on Earth differ from what we know today if no mutualistic relationships between fungi and other organisms had ever evolved?

For suggested answers, see Appendix A.

31 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 31.1

Fungi are heterotrophs that feed by absorption *(pp. 707–709)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

- All **fungi** (including decomposers and symbionts) are heterotrophs that acquire nutrients by absorption. Many fungi secrete enzymes that break down complex molecules.
- Most fungi grow as thin, multicellular filaments called **hyphae**; relatively few species grow only as single-celled **yeasts**. In their multicellular form, fungi consist of **mycelia**, networks of branched hyphae adapted for absorption. Mycorrhizal fungi have specialized hyphae that enable them to form a mutually beneficial relationship with plants.

? *How does the morphology of multicellular fungi affect the efficiency of nutrient absorption?*

CONCEPT 31.2

Fungi produce spores through sexual or asexual life cycles *(pp. 709–711)*

- In fungi, the sexual life cycle involves cytoplasmic fusion (**plasmogamy**) and nuclear fusion (**karyogamy**), with an intervening **heterokaryotic** stage in which cells have haploid nuclei from two parents. The diploid cells resulting from karyogamy are short-lived and undergo meiosis, producing genetically diverse haploid **spores**.
- Many fungi can reproduce asexually as filamentous fungi or yeasts.

DRAW IT ➤ *Draw a generalized fungal life cycle, labeling asexual and sexual reproduction, meiosis, plasmogamy, karyogamy, and the points in the cycle when spores and the zygote are produced.*

CONCEPT 31.3

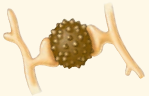
The ancestor of fungi was an aquatic, single-celled, flagellated protist *(pp. 711–712)*

- Molecular evidence indicates that fungi and animals diverged over a billion years ago from a common unicellular ancestor that had a flagellum. However, the oldest fossils that are widely accepted as fungi are 460 million years old.
- Chytrids, a group of fungi with flagellated spores, include some basal lineages.
- Fungi were among the earliest colonizers of land; fossil evidence indicates that these colonizers included species that were symbionts with early plants.

? *Did multicellularity originate independently in fungi and animals? Explain.*

CONCEPT 31.4

Fungi have radiated into a diverse set of lineages *(pp. 712–719)*

Fungal Phylum	Distinguishing Features
Chytridiomycota **(chytrids)**	Flagellated spores 
Zygomycota **(zygomycetes)**	Resistant zygosporangium as sexual stage 
Glomeromycota **(arbuscular mycorrhizal fungi)**	Arbuscular mycorrhizae formed with plants 
Ascomycota **(ascomycetes)**	Sexual spores (ascospores) borne internally in sacs called asci; vast numbers of asexual spores (conidia) produced 
Basidiomycota **(basidiomycetes)**	Elaborate fruiting body (basidiocarp) containing many basidia that produce sexual spores (basidiospores) 

DRAW IT ➤ *Draw a phylogenetic tree of the major groups of fungi.*

CONCEPT 31.5

Fungi play key roles in nutrient cycling, ecological interactions, and human welfare *(pp. 719–723)*

- Fungi perform essential recycling of chemical elements between the living and nonliving world.
- **Lichens** are highly integrated symbiotic associations of fungi and algae or cyanobacteria.
- Many fungi are parasites, mostly of plants.
- Humans use fungi for food and to make antibiotics.

? *How are fungi important as decomposers, mutualists, and pathogens?*

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–4 can be found in the Study Area in MasteringBiology.

5. SCIENTIFIC INQUIRY • INTERPRET THE DATA The grass *Dichanthelium lanuginosum* lives in hot soils and houses fungi of the genus *Curvularia* as endophytes. Researchers tested the impact of *Curvularia* on the heat tolerance of this grass. They grew plants without (E–) and with (E+) *Curvularia* endophytes at different temperatures and measured plant mass and the number of new shoots the plants produced. Draw a bar graph for plant mass versus temperature and interpret it.

PRACTICE TEST goo.gl/iAsVgL

Soil Temp.	*Curvularia* + or –	Plant Mass (g)	No. of New Shoots
30°C	E–	16.2	32
	E+	22.8	60
35°C	E–	21.7	43
	E+	28.4	60
40°C	E–	8.8	10
	E+	22.2	37
45°C	E–	0	0
	E+	15.1	24

Data from R. S. Redman et al., Thermotolerance generated by plant/fungal symbiosis, *Science* 298:1581 (2002).

6. EVOLUTION CONNECTION The fungus-alga symbiosis that makes up a lichen is thought to have evolved multiple times independently in different fungal groups. However, lichens fall into three well-defined growth forms (see Figure 31.22). How could you test the following hypotheses? Hypothesis 1: Crustose, foliose, and fruticose lichens each represent a monophyletic group. Hypothesis 2: Each lichen growth form represents convergent evolution by taxonomically diverse fungi.

7. WRITE ABOUT A THEME: ORGANIZATION Fungi are important decomposers. In a short essay (100–150 words), describe how the unique metabolic property of fungi benefits the energy cycle of ecosystems.

8. SYNTHESIZE YOUR KNOWLEDGE

This wasp is the unfortunate victim of an entomopathogenic fungus (a parasitic fungus of insects). Write a paragraph describing what this image illustrates about the nutritional mode, body structure, and ecological role of the fungus.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

An Introduction to Animal Diversity

32

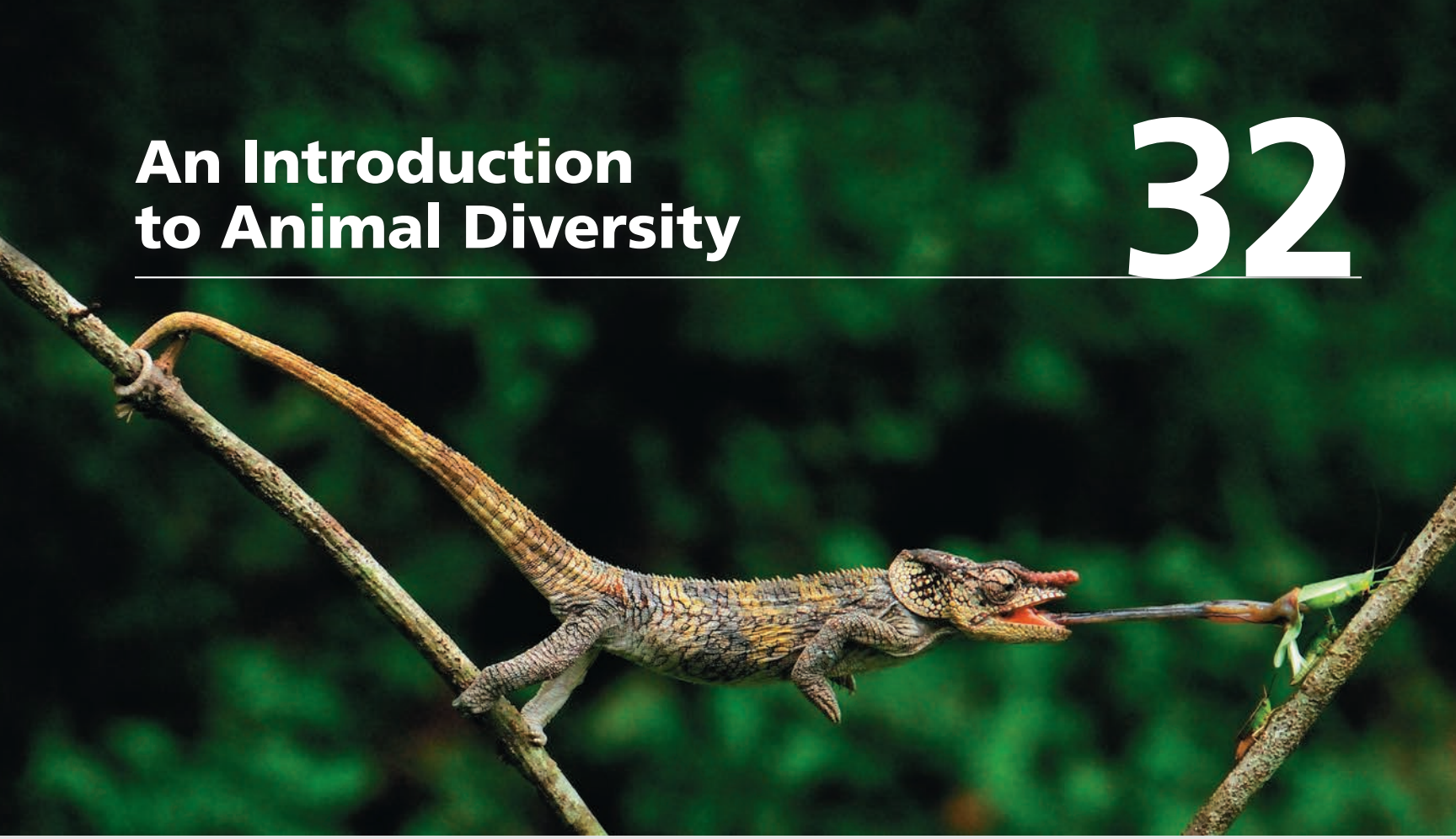


Figure 32.1 **What adaptations make a chameleon a fearsome predator?**

KEY CONCEPTS

32.1 Animals are multicellular, heterotrophic eukaryotes with tissues that develop from embryonic layers

32.2 The history of animals spans more than half a billion years

32.3 Animals can be characterized by "body plans"

32.4 Views of animal phylogeny continue to be shaped by new molecular and morphological data

A Kingdom of Consumers

Although slow-moving on its feet, the chameleon in **Figure 32.1** can wield its long, sticky tongue with blinding speed to capture its unsuspecting prey. Many species of chameleons can also change their color and thereby blend into their surroundings—making them hard to detect, both by their prey and by the animals that would eat them.

The chameleon is just one example of an animal that is an efficient consumer of other organisms. Other predatory animals overwhelm their prey using their strength, speed, or toxins, while still others capture the unwary by building concealed traps such as webs. Likewise, herbivorous animals can strip the plants they eat bare of leaves or seeds, while parasitic animals weaken their hosts by consuming their tissues or body fluids. These and other animals are effective eating machines in part because they have specialized muscle and nerve cells that enable them to detect, capture, and eat other organisms—including those that can flee from attack. Animals are also very good at processing the food they have eaten; most animals do this using an efficient digestive system that has a mouth at one end and an anus at the other.

In this chapter, we embark on a tour of the animal kingdom that will continue in the next two chapters. Here we will consider the characteristics that all animals share, as well as the evolutionary history of this kingdom of consumers.

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

CONCEPT 32.1

Animals are multicellular, heterotrophic eukaryotes with tissues that develop from embryonic layers

Listing features shared by all animals is challenging, as there are exceptions to nearly every criterion we might select. When taken together, however, several characteristics of animals sufficiently describe the group for our discussion.

Nutritional Mode

Animals differ from both plants and fungi in their mode of nutrition. Plants are autotrophic eukaryotes capable of generating organic molecules through photosynthesis. Fungi are heterotrophs that grow on or near their food and that feed by absorption (often after they have released enzymes that digest the food outside their bodies). Unlike plants, animals cannot construct all of their own organic molecules, and so, in most cases, they ingest them—either by eating other living organisms or by eating nonliving organic material. But unlike fungi, most animals feed by ingesting their food and then using enzymes to digest it within their bodies.

Cell Structure and Specialization

Animals are eukaryotes, and like plants and most fungi, animals are multicellular. In contrast to plants and fungi, however, animals lack the structural support of cell walls. Instead, proteins external to the cell membrane provide structural support to animal cells and connect them to one another (see Figure 7.28). The most abundant of these proteins is collagen, which is not found in plants or fungi.

The cells of most animals are organized into **tissues**, groups of similar cells that act as a functional unit. For example, muscle tissue and nervous tissue are responsible for moving the body and conducting nerve impulses, respectively. The ability to move and conduct nerve impulses underlies many of the adaptations that differentiate animals from plants and fungi (which lack muscle and nerve cells). For this reason, muscle and nerve cells are central to the animal lifestyle.

Reproduction and Development

Most animals reproduce sexually, and the diploid stage usually dominates the life cycle. In the haploid stage, sperm and egg cells are produced directly by meiotic division, unlike what occurs in plants and fungi (see Figure 13.6). In most animal species, a small, flagellated sperm fertilizes a larger, nonmotile egg, forming a diploid zygote. The zygote then undergoes **cleavage**, a succession of mitotic cell divisions without cell growth between the divisions. During the development of most animals, cleavage leads to the formation of a multicellular embryonic stage called a **blastula**, which in many animals takes the form of a hollow ball **(Figure 32.2)**. Following this stage is the process of **gastrulation**, during which the layers of embryonic tissues that will develop into adult body parts are produced (see also Figure 46.8). The resulting developmental stage is called a **gastrula**.

Although some animals, including humans, develop directly into adults, the life cycles of most animals include at least one larval stage. A **larva** is a sexually immature form of an animal that is morphologically distinct from the adult, usually eats different food, and may even have a different habitat than the adult, as in the case of the aquatic larva of a mosquito or dragonfly. Animal larvae eventually undergo **metamorphosis**, a developmental transformation that turns the animal into a juvenile that resembles an adult but is not yet sexually mature.

Though adult animals vary widely in morphology, the genes that control animal development are similar across a broad range of taxa.

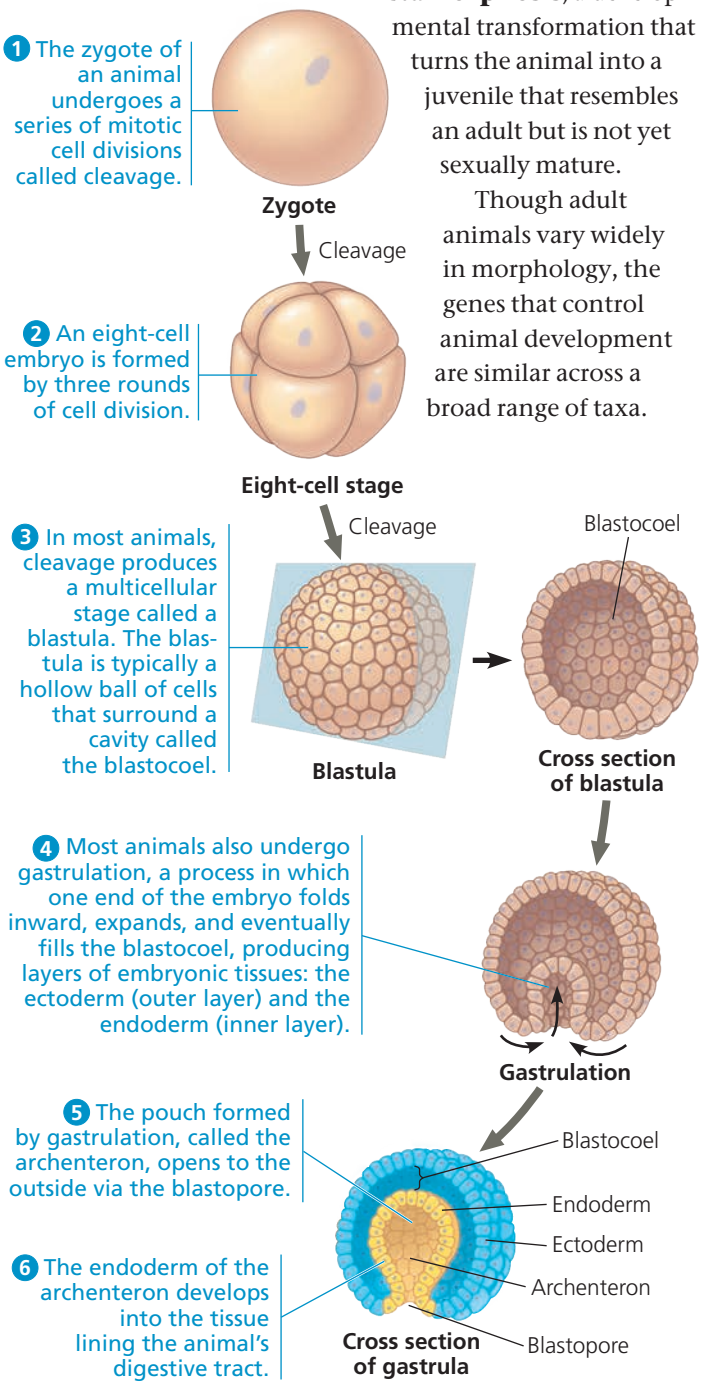


Figure 32.2 Early embryonic development in animals.

All animals have developmental genes that regulate the expression of other genes, and many of these regulatory genes contain sets of DNA sequences called *homeoboxes* (see Concept 20.6). In particular, most animals share a unique homeobox-containing family of genes, known as *Hox* genes. *Hox* genes play important roles in the development of animal embryos, controlling the expression of many other genes that influence morphology.

Sponges, which are among the simplest extant (living) animals, lack *Hox* genes. However, they have other homeobox genes that influence their shape, such as those that regulate the formation of water channels in the body wall, a key feature of sponge morphology (see Figure 33.4). In the ancestors of more complex animals, the *Hox* gene family arose via the duplication of earlier homeobox genes. Over time, the *Hox* gene family underwent a series of duplications, yielding a versatile "toolkit" for regulating development. In most animals, *Hox* genes regulate the formation of the anterior-posterior (front-to-back) axis, as well as other aspects of development. Similar sets of conserved genes govern the development of both flies and humans, despite their obvious differences and hundreds of millions of years of divergent evolution.

CONCEPT CHECK 32.1

1. Summarize the main stages of animal development. What family of control genes plays a major role?
2. **WHAT IF?** ➤ What animal characteristics would be needed by an imaginary plant that could chase, capture, and digest its prey—yet could also extract nutrients from soil and conduct photosynthesis?

For suggested answers, see Appendix A.

CONCEPT 32.2

The history of animals spans more than half a billion years

To date, biologists have identified 1.3 million extant species of animals, and estimates of the actual number run far higher. This vast diversity encompasses a spectacular range of morphological variation, from corals to cockroaches to crocodiles. Various studies suggest that this great diversity originated during the last billion years. For example, researchers have unearthed 710-million-year-old sediments containing chemical evidence of steroids that today are primarily produced by a particular group of sponges. Since sponges are animals, these "fossil steroids" suggest that animals had arisen by 710 million years ago.

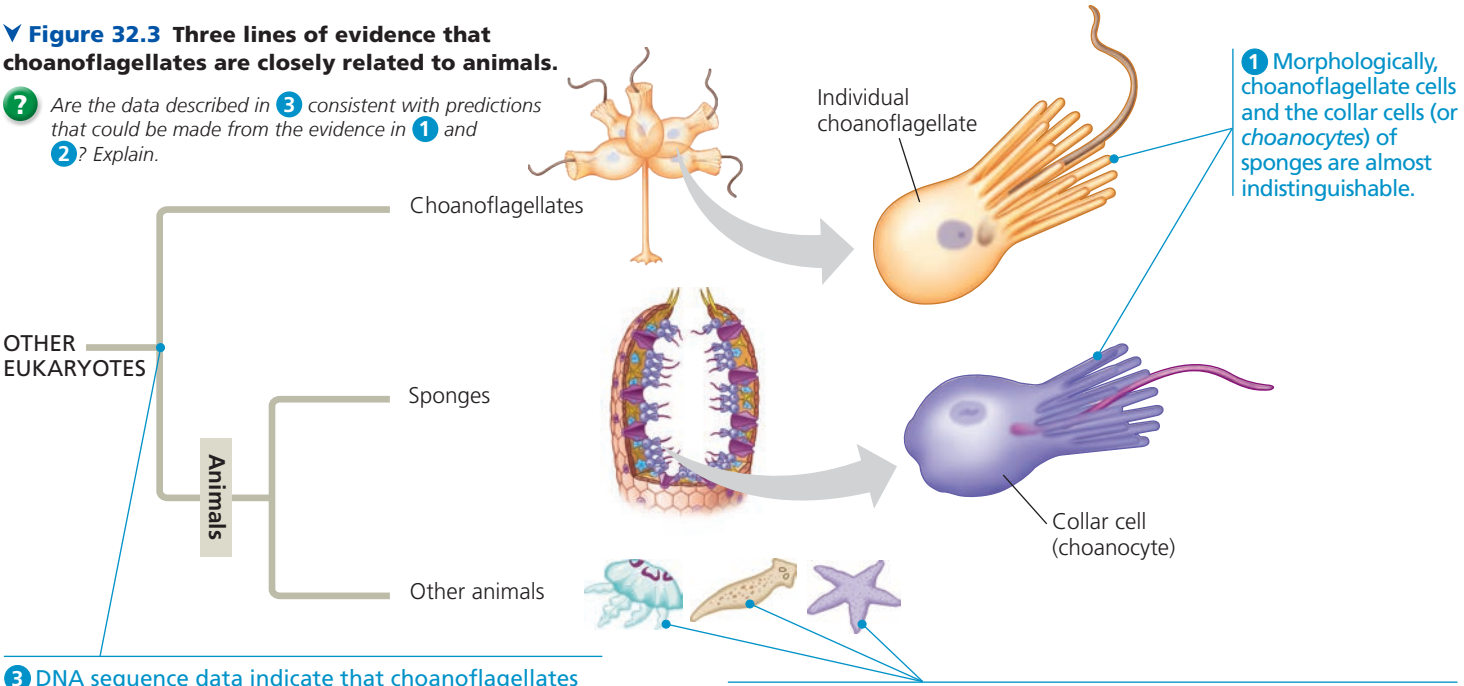
DNA analyses generally agree with this fossil biochemical evidence; for example, one recent molecular clock study estimated that sponges originated about 700 million years ago. These findings are also consistent with molecular analyses suggesting that the common ancestor of all extant animal species lived about 770 million years ago. What was this common ancestor like, and how did animals arise from their single-celled ancestors?

Steps in the Origin of Multicellular Animals

One way to gather information about the origin of animals is to identify protist groups that are closely related to animals. As shown in **Figure 32.3**, a combination of morphological and molecular evidence points to choanoflagellates as the

Figure 32.3 Three lines of evidence that choanoflagellates are closely related to animals.

? *Are the data described in* 3 *consistent with predictions that could be made from the evidence in* 1 *and* 2*? Explain.*



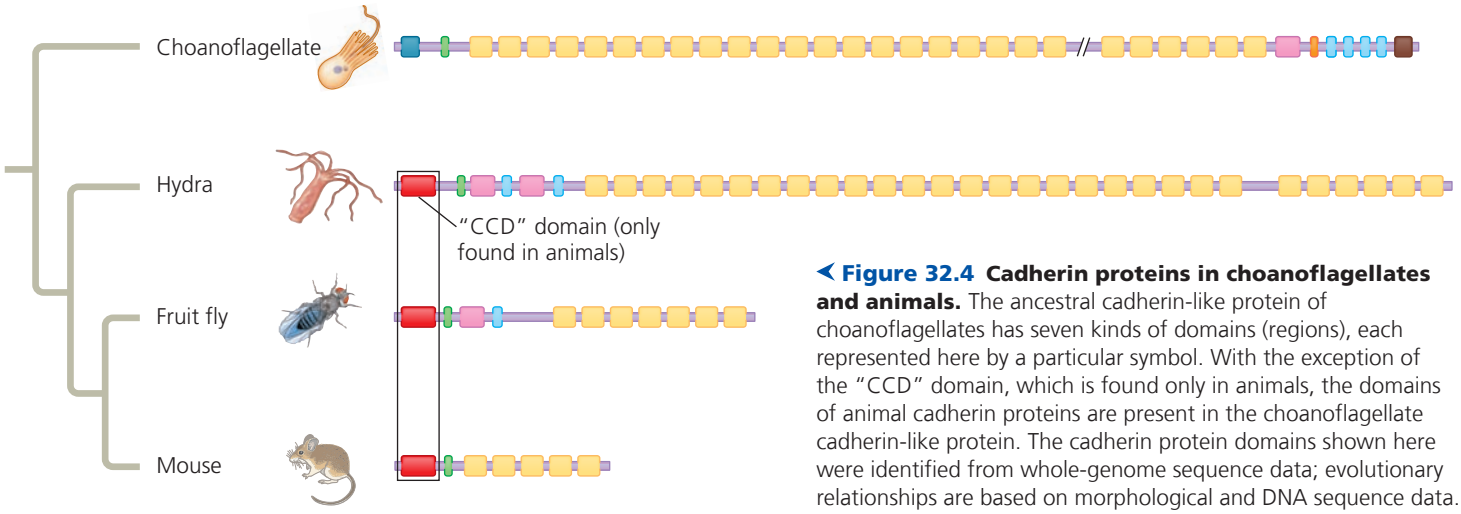


Figure 32.4 Cadherin proteins in choanoflagellates and animals. The ancestral cadherin-like protein of choanoflagellates has seven kinds of domains (regions), each represented here by a particular symbol. With the exception of the "CCD" domain, which is found only in animals, the domains of animal cadherin proteins are present in the choanoflagellate cadherin-like protein. The cadherin protein domains shown here were identified from whole-genome sequence data; evolutionary relationships are based on morphological and DNA sequence data.

closest living relatives of animals. Based on such evidence, researchers have hypothesized that the common ancestor of choanoflagellates and living animals may have been a suspension feeder similar to present-day choanoflagellates.

Scientists exploring *how* animals may have arisen from their single-celled ancestors have noted that the origin of multicellularity requires the evolution of new ways for cells to adhere (attach) and signal (communicate) to each other. In an effort to learn more about such mechanisms, researchers compared the genome of the unicellular choanoflagellate *Monosiga brevicollis* with those of representative animals. This analysis uncovered 78 protein domains in *M. brevicollis* that were otherwise only known to occur in animals. (A *domain* is a key structural or functional region of a protein.) For example, *M. brevicollis* has genes that encode domains of certain proteins (known as cadherins) that play key roles in how animal cells attach to one another, as well as genes that encode protein domains that animals (and only animals) use in cell-signaling pathways.

Let's take a closer look at the cadherin attachment proteins we just mentioned. DNA sequence analyses show that animal cadherin proteins are composed primarily of domains that are also found in a cadherin-like protein of choanoflagellates **(Figure 32.4)**. However, animal cadherin proteins also contain a highly conserved region not found in the choanoflagellate protein (the "CCD" domain labeled in Figure 32.4). These data suggest that the cadherin attachment protein originated by the rearrangement of protein domains found in choanoflagellates plus the incorporation of a novel domain, the conserved CCD region. Overall, comparisons of choanoflagellate and animal genomes suggest that key steps in the transition to multicellularity in animals involved new ways of using proteins or parts of proteins that were encoded by genes found in choanoflagellates.

Next, we'll survey the fossil evidence for how animals evolved from their distant common ancestor over four geologic eras (see Table 25.1 to review the geologic time scale).

Neoproterozoic Era (1 Billion–541 Million Years Ago)

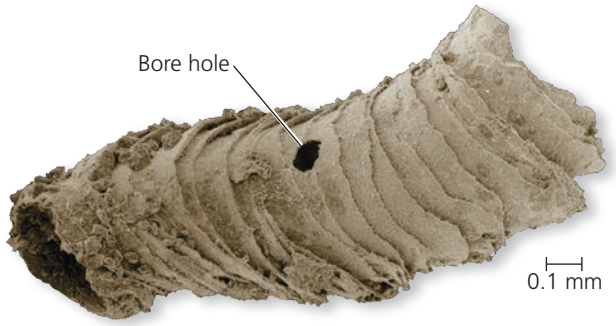
Although data from fossil steroids and molecular clocks indicate an earlier origin, the first generally accepted macroscopic fossils of animals date from about 560 million years ago. These fossils are members of an early group of soft-bodied multicellular eukaryotes, known collectively as the **Ediacaran biota**. The name comes from the Ediacara Hills of Australia, where fossils of these organisms were first discovered **(Figure 32.5)**. Similar fossils have since been found on other continents. Among the oldest Ediacaran fossils that resemble animals, some have been classified as molluscs (snails and their relatives) or close relatives of the molluscs, while others are thought to be sponges or cnidarians (sea anemones and their relatives). Still others have proved difficult to classify, as they do not seem to be closely related to any living animal or algal groups. In addition to these macroscopic fossils, Neoproterozoic rocks have also yielded what may be

(a) *Dickinsonia costata* (taxonomic affiliation unknown)

(b) *Kimberella*, a mollusc (or close relative)

Figure 32.5 Ediacaran fossil animals. Fossils dating to 560 million years ago document the earliest known macroscopic animals, including these two species.

Figure 32.6 Early evidence of predation. This 550-million-year-old fossil of the animal *Cloudina* shows evidence of having been attacked by a predator that bored through its shell.



microscopic fossils of early animal embryos. Although these microfossils appear to exhibit the basic structural organization of present-day animal embryos, debate continues about whether these fossils are indeed of animals.

The fossil record from the Ediacaran period (635–541 million years ago) also provides early evidence of predation. Consider *Cloudina*, a small animal whose body was protected by a shell resembling a series of nested cones **(Figure 32.6)**. Some *Cloudina* fossils show signs of attack: round "bore holes" that resemble those formed today by predators that drill through the shells of their prey to gain access to the soft-bodied organisms lying within. Like *Cloudina*, some other small Ediacaran animals had shells or other defensive structures that may have been selected for by predators. Overall, the fossil evidence indicates that the Ediacaran was a time of increasing animal diversity—a trend that continued in the Paleozoic.

Paleozoic Era (541–252 Million Years Ago)

Another wave of animal diversification occurred 535–525 million years ago, during the Cambrian period of the Paleozoic era—a phenomenon referred to as the **Cambrian explosion** (see Concept 25.3). In strata formed before the Cambrian explosion, only a few animal phyla have been observed. But in strata that are 535–525 million years old, paleontologists have found the oldest fossils of about half of all extant animal phyla, including the first arthropods, chordates, and echinoderms. Many of these fossils, which include the first large animals with hard, mineralized skeletons, look very different from most living animals **(Figure 32.7)**. Even so, paleontologists have established that these Cambrian fossils are members of extant animal phyla, or at least are close relatives. In particular, most of the fossils from the Cambrian explosion are of **bilaterians**, an enormous clade whose members (unlike sponges and cnidarians) typically have a two-sided or bilaterally symmetric form and a complete digestive tract, an efficient digestive system that has a mouth at one end and an anus at the other. As we'll discuss later in the chapter, bilaterians include molluscs, arthropods, chordates, and most other living animal phyla.

As the diversity of animal phyla increased during the Cambrian, the diversity of Ediacaran life-forms declined. What caused these trends? Fossil evidence suggests that during the Cambrian period, predators acquired novel adaptations, such as forms of locomotion that helped them catch prey, while prey species acquired new defenses, such as protective shells. As new predator-prey relationships emerged, natural selection may have led to the decline of the soft-bodied Ediacaran species and the rise of various bilaterian phyla. Another hypothesis focuses on an increase in atmospheric oxygen that preceded the Cambrian explosion. More plentiful oxygen would have enabled animals with higher metabolic rates and larger body sizes to thrive, while potentially harming other species. A third hypothesis proposes that genetic changes affecting development, such as the origin of *Hox* genes and the addition of new microRNAs (small RNAs involved in gene regulation), facilitated the evolution of new body forms. In the **Scientific Skills Exercise**, you can investigate whether there is a correlation between microRNAs (miRNAs; see Figure 18.14) and body complexity in various animal phyla. These various hypotheses are not mutually exclusive; predator-prey relationships, atmospheric changes, and changes in development may each have played a role.

The Cambrian period was followed by the Ordovician, Silurian, and Devonian periods, when animal diversity

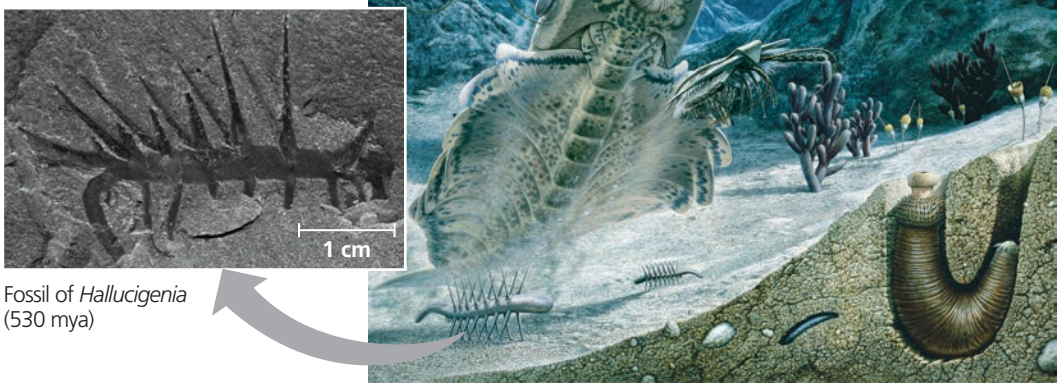


Figure 32.7 A Cambrian seascape. This artist's reconstruction depicts a diverse array of organisms found in fossils from the Burgess Shale site in British Columbia, Canada. The animals include *Pikaia* (eel-like chordate at top left), *Marella* (small arthropod swimming at left), *Anomalocaris* (large animal with grasping limbs and a circular mouth), and *Hallucigenia* (animals with toothpick-like spikes on the seafloor and in inset).

SCIENTIFIC SKILLS EXERCISE

Calculating and Interpreting Correlation Coefficients

Is Animal Complexity Correlated with miRNA Diversity? Animal phyla vary greatly in morphology, from simple sponges that lack tissues and symmetry to complex vertebrates. Members of different animal phyla have similar developmental genes, but the number of miRNAs varies considerably. In this exercise, you will explore whether miRNA diversity is correlated to morphological complexity.

How the Study Was Done In the analysis, miRNA diversity is represented by the average number of miRNAs in a phylum (x), while morphological complexity is represented by the average number of cell types (y). Researchers examined the relationship between these two variables by calculating the correlation coefficient (r). The correlation coefficient indicates the extent and direction of a linear relationship between two variables (x and y) and ranges in value between −1 and 1. When $r < 0$, y and x are negatively correlated, meaning that values of y become smaller as values of x become larger. When $r > 0$, y and x are positively correlated (y becomes larger as x becomes larger). When $r = 0$, the variables are not correlated.

The formula for the correlation coefficient r is the following:

$$r = \frac{\frac{1}{n-1}\sum (x_i - \bar{x})(y_i - \bar{y})}{s_x s_y}$$

In this formula, n is the number of observations, x_i is the value of the i^{th} observation of variable x, and y_i is the value of the i^{th} observation of variable y. $\bar{x}$ and $\bar{y}$ are the means of variables x and y, and s_x and s_y are the standard deviations of variables x and y. The "Σ" symbol indicates that the n values of the product $(x_i - \bar{x})(y_i - \bar{y})$ are to be added together.

Data from the Study

Animal Phylum	i	No. of miRNAs (x_i)	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	No. of Cell Types (y_i)	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
Porifera	1	5.8			25			
Platyhelminthes	2	35			30			
Cnidaria	3	2.5			34			
Nematoda	4	26			38			
Echinodermata	5	38.6			45			
Cephalochordata	6	33			68			
Arthropoda	7	59.1			73			
Urochordata	8	25			77			
Mollusca	9	50.8			83			
Annelida	10	58			94			
Vertebrata	11	147.5			172.5			
		$\bar{x}$ = s_x =		Σ =	$\bar{y}$ = s_y =		Σ =	Σ =

Data from Bradley Deline, University of West Georgia, and Kevin Peterson, Dartmouth College, 2013.

INTERPRET THE DATA

1. First, practice reading the data table. For the eighth observation ($i = 8$), what are x_i and y_i? For which phylum are these data?
2. Next, we'll calculate the mean and standard deviation for each variable. (a) The **mean** ($\bar{x}$) is the sum of the data values divided by n, the number of observations: $\bar{x} = \frac{\Sigma x_i}{n}$. Calculate the mean number of miRNAs ($\bar{x}$) and the mean number of cell types ($\bar{y}$) and enter them in the data table (for $\bar{y}$, replace each x in the formula with a y). (b) Next, calculate $(x_i - \bar{x})$ and $(y_i - \bar{y})$ for each observation, recording your results in the appropriate column. Square each of those results to complete the $(x_i - \bar{x})^2$ and $(y_i - \bar{y})^2$ columns; sum the results for those columns. (c) The **standard deviation**, s_x, which describes the variation found in the data, is calculated using the following formula:

$$s_x = \sqrt{\frac{1}{n-1}\sum (x_i - \bar{x})^2}$$

   Calculate s_x and s_y by substituting the results in (b) into the formula for the standard deviation.
3. Next, calculate the correlation coefficient r for the variables x and y. (a) First, use the results in 3(b) to complete the $(x_i - \bar{x})(y_i - \bar{y})$ column; sum the results in that column. (b) Now use the values for s_x and s_y from 3(c) along with the results from 4(a) in the formula for r.
4. Do these data indicate that miRNA diversity and animal complexity are negatively correlated, positively correlated, or uncorrelated? Explain.
5. What does your analysis suggest about the role of miRNA diversity in the evolution of animal complexity?

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

continued to increase, although punctuated by episodes of mass extinction (see Figure 25.17). Vertebrates (fishes) emerged as the top predators of the marine food web. By 450 million years ago, groups that diversified during the Cambrian period began to make an impact on land. Arthropods were the first animals to adapt to terrestrial habitats, as indicated by fragments of arthropod remains and by well-preserved fossils from several continents of millipedes, centipedes, and spiders. Another clue is seen in fossilized fern galls—enlarged cavities that fern plants form in response to stimulation by resident insects, which then use the galls for protection. Fossils indicate that fern galls date back at least 302 million years, suggesting that insects and plants were influencing each other's evolution by that time.

Vertebrates colonized land around 365 million years ago and diversified into numerous terrestrial groups. Two of these survive today: the amphibians (such as frogs and salamanders) and the amniotes (reptiles, including birds, and mammals). We will explore these groups, known collectively as the tetrapods, in more detail in Chapter 34.

Mesozoic Era (252–66 Million Years Ago)

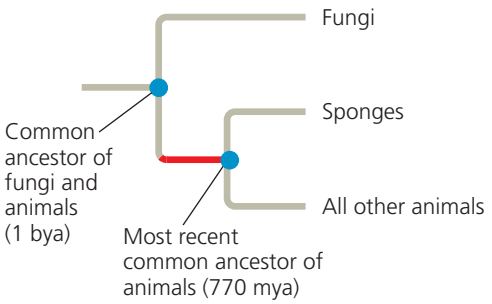
The animal phyla that had evolved during the Paleozoic now began to spread into new habitats. In the oceans, the first coral reefs formed, providing other marine animals with new places to live. Some reptiles returned to the water, leaving plesiosaurs (see Figure 25.5) and other large aquatic predators as their descendants. On land, descent with modification in some tetrapods led to the origin of wings and other flight equipment in pterosaurs and birds. Large and small dinosaurs emerged, both as predators and herbivores. At the same time, the first mammals—tiny nocturnal insect-eaters—appeared on the scene. In addition, as you read in Concept 30.3, flowering plants (angiosperms) and insects both underwent dramatic diversifications during the late Mesozoic.

Cenozoic Era (66 Million Years Ago to the Present)

Mass extinctions of both terrestrial and marine animals ushered in a new era, the Cenozoic. Among the groups of species that disappeared were the large, nonflying dinosaurs and the marine reptiles. The fossil record of the early Cenozoic documents the rise of large mammalian herbivores and predators as mammals began to exploit the vacated ecological niches. The global climate gradually cooled throughout the Cenozoic, triggering significant shifts in many animal lineages. Among primates, for example, some species in Africa adapted to the open woodlands and savannas that replaced many of the former dense forests. The ancestors of our own species were among those grassland apes.

CONCEPT CHECK 32.2

1. Put the following milestones in animal evolution in order from oldest to most recent: (a) origin of mammals, (b) earliest evidence of terrestrial arthropods, (c) Ediacaran fauna, (d) extinction of large, nonflying dinosaurs.
2. **VISUAL SKILLS ➤** Explain what is represented by the red-colored portion of the branch leading to animals. (See Figure 22.5, "Visualizing Phylogenetic Relationships," to review phylogenetic tree diagrams.)



3. **MAKE CONNECTIONS ➤** Evaluate whether the origin of cell-to-cell attachment proteins in animals illustrates descent with modification. (See Concept 21.2.)

For suggested answers, see Appendix A.

CONCEPT 32.3

Animals can be characterized by "body plans"

Animal species vary tremendously in morphology, but their great diversity in form can be described by a relatively small number of major "body plans." A **body plan** is a particular set of morphological and developmental traits, integrated into a functional whole—the living animal. The term *plan* here does not imply that animal forms are the result of conscious planning or invention. But body plans do provide a succinct way to compare and contrast key animal features. They also are of interest in the study of *evo-devo*, the interface between evolution and development.

Like all features of organisms, animal body plans have evolved over time. In some cases, including key stages in gastrulation, novel body plans emerged early in the history of animal life and have not changed since. As we'll discuss, however, other aspects of animal body plans have changed multiple times over the course of evolution. As we explore the major features of animal body plans, bear in mind that similar body forms may have evolved independently in different lineages. In addition, body features can be lost over the course of evolution, causing some closely related species to look very different from one another.

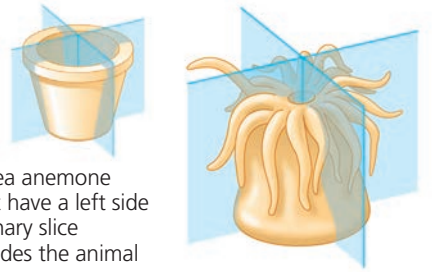
Symmetry

A basic feature of animal bodies is their type of symmetry—or absence of symmetry. (Many sponges, for example, lack symmetry altogether.) Some animals exhibit **radial symmetry**, the type of symmetry found in a flowerpot **(Figure 32.8a)**. Sea anemones, for example, have a top side (where the mouth is located) and a bottom side. But they have no front and back ends and no left and right sides.

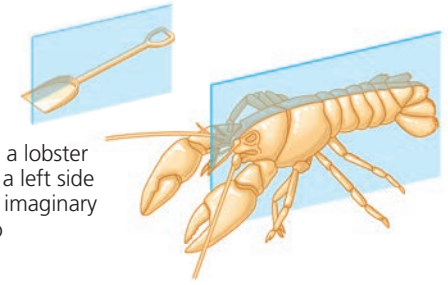
The two-sided symmetry of a shovel is an example of **bilateral symmetry (Figure 32.8b)**. A bilateral animal has two axes of orientation: front to back and top to bottom. Such animals have a **dorsal** (top) side and a **ventral** (bottom) side, a left side and a right side, and an **anterior** (front) end and a **posterior** (back) end. Nearly all animals with a bilaterally symmetrical body plan (such as arthropods and mammals) have sensory equipment concentrated at their anterior end, including a central nervous system ("brain") in the head.

The symmetry of an animal generally fits its lifestyle. Many radial animals are sessile (living attached to a substrate) or planktonic (drifting or weakly swimming, such as jellies, commonly called jellyfishes). Their symmetry equips them to meet the environment equally well from all sides. In contrast, bilateral animals typically move actively from place to place. Most bilateral animals have a central nervous system that enables them to coordinate the complex movements involved in crawling, burrowing, flying,

Figure 32.8 Body symmetry. The flowerpot and shovel are included to help you remember the radial-bilateral distinction.



(a) Radial symmetry. A radial animal, such as a sea anemone (phylum Cnidaria), does not have a left side and a right side. Any imaginary slice through the central axis divides the animal into mirror images.



(b) Bilateral symmetry. A bilateral animal, such as a lobster (phylum Arthropoda), has a left side and a right side. Only one imaginary cut divides the animal into mirror-image halves.

or swimming. Fossil evidence indicates that these two fundamentally different kinds of symmetry have existed for at least 550 million years.

Tissues

Animal body plans also vary with regard to tissue organization. Recall that tissues are collections of specialized cells that act as a functional unit. Sponges and a few other groups lack tissues. In all other animals, the embryo becomes layered during gastrulation (see Figure 46.8, "Visualizing Gastrulation," which will help you understand this three-dimensional folding process). As development progresses, these layers, called *germ layers*, form the various tissues and organs of the body. **Ectoderm**, the germ layer covering the surface of the embryo, gives rise to the outer covering of the animal and, in some phyla, to the central nervous system. **Endoderm**, the innermost germ layer, lines the pouch that forms during gastrulation (the archenteron) and gives rise to the lining of the digestive tract (or cavity) and to the lining of organs such as the liver and lungs of vertebrates.

Cnidarians and a few other animal groups that have only these two germ layers are said to be **diploblastic**. All bilaterally symmetrical animals have a third germ layer, called the **mesoderm**, which fills much of the space between the ectoderm and endoderm. Thus, animals with bilateral symmetry are also said to be **triploblastic** (having three germ layers). In triploblasts, the mesoderm forms the muscles and most other organs between the digestive tract and the outer covering of the animal. Triploblasts include a broad range of animals, from flatworms to arthropods to vertebrates. (Although some diploblasts actually do have a third germ layer, it is not nearly as well developed as the mesoderm of animals considered to be triploblastic.)

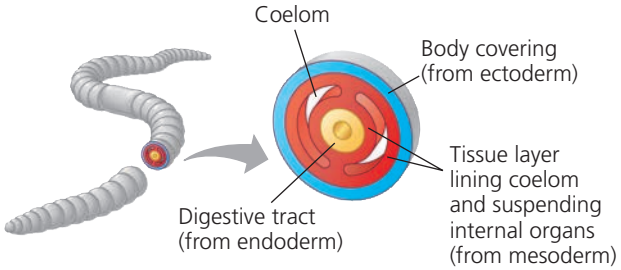
Body Cavities

Most triploblastic animals have a **body cavity**, a fluid- or air-filled space located between the digestive tract and the outer body wall. This body cavity is also called a **coelom** (from the Greek *koilos*, hollow). A so-called "true" coelom forms from tissue derived from mesoderm. The inner and outer layers of tissue that surround the cavity connect and form structures that suspend the internal organs. Animals with a true coelom are known as **coelomates** **(Figure 32.9a)**.

Some triploblastic animals have a body cavity that is formed from mesoderm and endoderm **(Figure 32.9b)**. Such a cavity is called a "pseudocoelom" (from the Greek

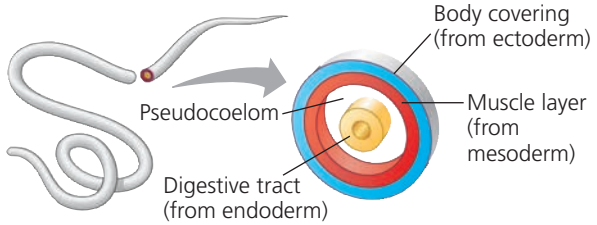
Figure 32.9 Body cavities of triploblastic animals. The organ systems develop from the three embryonic germ layers.

(a) Coelomate



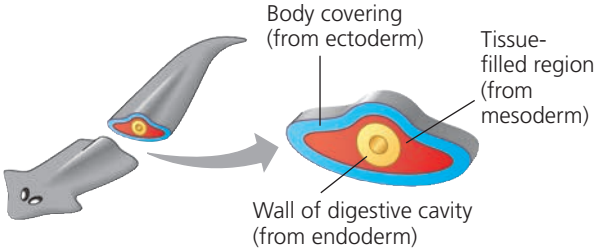
Coelomates, such as earthworms, have a true coelom, a body cavity completely lined by tissue derived from mesoderm.

(b) Pseudocoelomate



Pseudocoelomates, such as roundworms, have a body cavity lined by tissue derived from mesoderm and by tissue derived from endoderm.

(c) Acoelomate



Acoelomates, such as planarians, lack a body cavity between the digestive cavity and outer body wall.

Key

Ectoderm Mesoderm Endoderm

pseudo, false), and the animals that have one are called **pseudocoelomates**. Despite its name, however, a pseudocoelom is not false; it is a fully functional body cavity. Finally, some triploblastic animals lack a body cavity altogether **(Figure 32.9c)**. They are known collectively as **acoelomates** (from the Greek *a-*, without).

A body cavity has many functions. Its fluid cushions the suspended organs, helping to prevent internal injury. In soft-bodied coelomates, such as earthworms, the coelom contains noncompressible fluid that acts like a skeleton against which muscles can work. The cavity also enables the internal organs to grow and move independently of the outer body wall. If it were not for your coelom, for example, every beat of your heart or ripple of your intestine would warp your body's surface.

Terms such as *coelomates* and *pseudocoelomates* refer to organisms that have a similar body plan and hence belong to the same *grade* (a group whose members share key biological features). However, phylogenetic studies show that true coeloms and pseudocoeloms have been independently gained or lost multiple times in the course of animal evolution. As shown by this example, a grade is not necessarily equivalent to a *clade* (a group that includes an ancestral species and all of its descendants). Thus, while terms such as coelomate or pseudocoelomate can be helpful in describing an organism's features, these terms must be interpreted with caution when seeking to understand evolutionary history.

Protostome and Deuterostome Development

Based on certain aspects of early development, many animals can be described as having one of two developmental modes: **protostome development** or **deuterostome development**. These modes can generally be distinguished by differences in cleavage, coelom formation, and fate of the blastopore.

Cleavage

Many animals with protostome development undergo **spiral cleavage**, in which the planes of cell division are diagonal to the vertical axis of the embryo; as seen in the eight-cell stage of the embryo, smaller cells are centered over the grooves between larger, underlying cells **(Figure 32.10a**, left). Furthermore, the so-called **determinate cleavage** of some animals with protostome development rigidly casts ("determines") the developmental fate of each embryonic cell very early. A cell isolated from a snail at the four-cell stage, for example, cannot develop into a whole animal. Instead, after repeated divisions, such a cell will form an inviable embryo that lacks many parts.

In contrast to the spiral cleavage pattern, deuterostome development is predominantly characterized by **radial cleavage**. The cleavage planes are either parallel or perpendicular to the vertical axis of the embryo; as seen at the

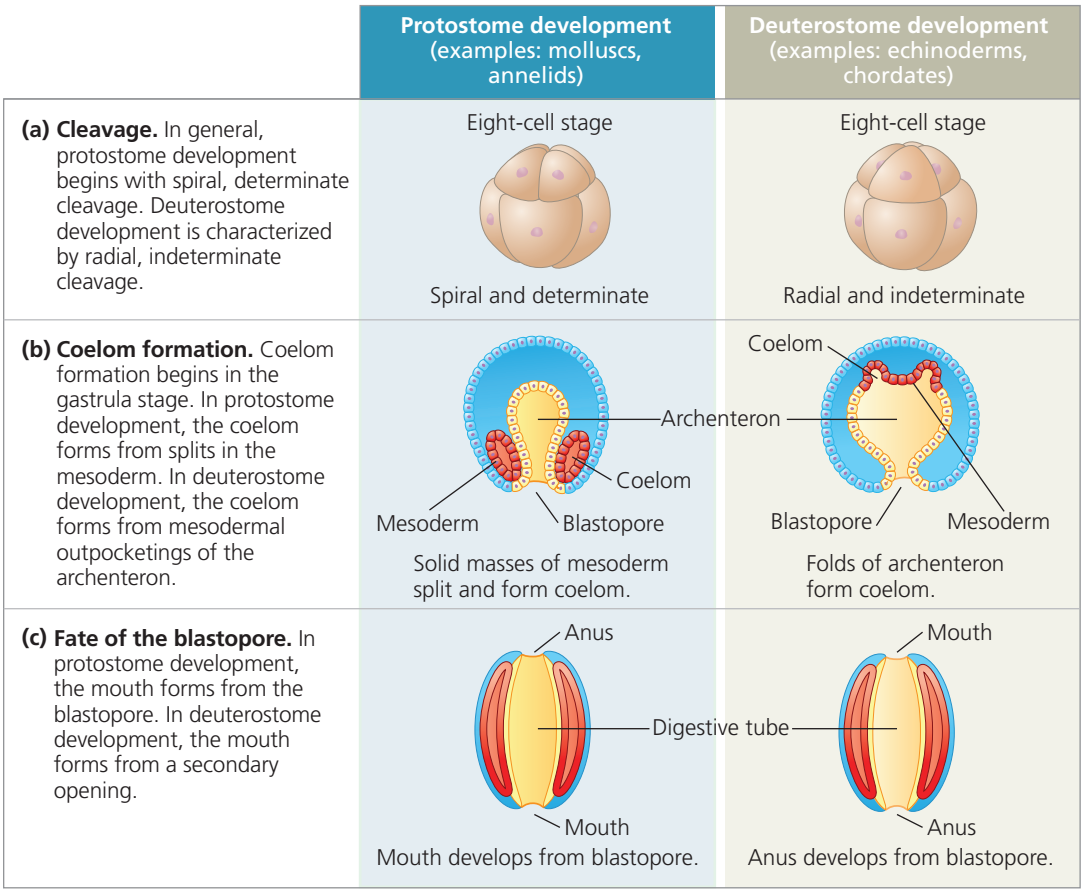


Figure 32.10 A comparison of protostome and deuterostome development. These are useful general distinctions, though there are many variations and exceptions to these patterns.

MAKE CONNECTIONS *Review Figure 19.20. As an early embryo, which of the following would more likely have stem cells capable of giving rise to cells of any type: an animal with protostome development or one with deuterostome development? Explain.*

eight-cell stage, the tiers of cells are aligned, one directly above the other (see Figure 32.10a, right). Most animals with deuterostome development also have **indeterminate cleavage**, meaning that each cell produced by early cleavage divisions retains the capacity to develop into a complete embryo. For example, if the cells of a sea urchin embryo are separated at the four-cell stage, each can form a complete larva. Similarly, it is the indeterminate cleavage of the human zygote that makes identical twins possible.

Coelom Formation

During gastrulation, an embryo's developing digestive tube initially forms as a blind pouch, the **archenteron**, which becomes the gut **(Figure 32.10b)**. As the archenteron forms in protostome development, initially solid masses of mesoderm split and form the coelom. In contrast, in deuterostome development, the mesoderm buds from the wall of the archenteron, and its cavity becomes the coelom.

Fate of the Blastopore

Protostome and deuterostome development often differ in the fate of the **blastopore**, the indentation that during gastrulation leads to the formation of the archenteron **(Figure 32.10c)**. After the archenteron develops, in most animals a second opening forms at the opposite end of the gastrula. In many species, the blastopore and this second opening become the two openings of the digestive tube: the mouth and the anus. In protostome development, the mouth generally develops from the first opening, the blastopore, and it is for this characteristic that the term *protostome* derives (from the Greek *protos*, first, and *stoma*, mouth). In deuterostome development (from the Greek *deuteros*, second), the mouth is derived from the secondary opening, and the blastopore usually forms the anus.

CONCEPT CHECK 32.3

1. Distinguish the terms *grade* and *clade*.
2. Name an organ or body part formed from each of the germ layers in a triploblastic animal.
3. **WHAT IF?** ➤ Evaluate this claim: Ignoring the details of their specific anatomy, worms, humans, and most other triploblasts have a shape analogous to that of a doughnut.

For suggested answers, see Appendix A.

CONCEPT 32.4

Views of animal phylogeny continue to be shaped by new molecular and morphological data

As animals with diverse body plans radiated during the early Cambrian, some lineages arose, thrived for a period of time, and then became extinct, leaving no descendants. However, by 500 million years ago, most animal phyla with members alive today were established. Next, we'll examine relationships among these taxa along with some remaining questions that are currently being addressed using genomic data.

The Diversification of Animals

Zoologists currently recognize about three dozen phyla of extant animals, 15 of which are shown in **Figure 32.11**. Researchers infer evolutionary relationships among these phyla by analyzing whole genomes, as well as morphological traits, ribosomal RNA (rRNA) genes, *Hox* genes, protein-coding nuclear genes, and mitochondrial genes. Notice how the following points are reflected in Figure 32.11.

1. **All animals share a common ancestor.** Current evidence indicates that animals are monophyletic, forming a clade called Metazoa. All extant and extinct animal lineages have descended from a common ancestor.
2. **Sponges are the sister group to all other animals.** Sponges (phylum Porifera) are basal animals, having diverged from all other animals early in the history of the group. Recent morphological and molecular analyses indicate that sponges are monophyletic, as shown here.
3. **Eumetazoa is a clade of animals with tissues.** All animals except for sponges and a few others belong to a clade of **eumetazoans** ("true animals"). Members of this group have tissues, such as muscle tissue and nervous tissue. Basal eumetazoans, which include the phyla Ctenophora (comb jellies) and Cnidaria, are diploblastic and generally have radial symmetry.
4. **Most animal phyla belong to the clade Bilateria.** Bilateral symmetry and the presence of three prominent germ layers are shared derived characters that help define the clade Bilateria. This clade contains the majority of animal phyla, and its members are known as *bilaterians*. The Cambrian explosion was primarily a rapid diversification of bilaterians.
5. **There are three major clades of bilaterian animals.** Bilaterians have diversified into three main lineages, Deuterostomia, Lophotrochozoa, and Ecdysozoa. With one exception, the phyla in these clades consist entirely of **invertebrates**, animals that lack a backbone; Chordata is the only phylum that includes **vertebrates**, animals with a backbone.

As seen in Figure 32.11, hemichordates (acorn worms), echinoderms (sea stars and relatives), and chordates are members of the bilaterian clade **Deuterostomia**; thus, the term *deuterostome* refers not only to a mode of animal development, but also to the members of this clade. (The dual meaning of this term can be confusing since some organisms with a deuterostome developmental pattern are *not* members of clade Deuterostomia.) Hemichordates share some characteristics with chordates, such as gill slits and a dorsal nerve

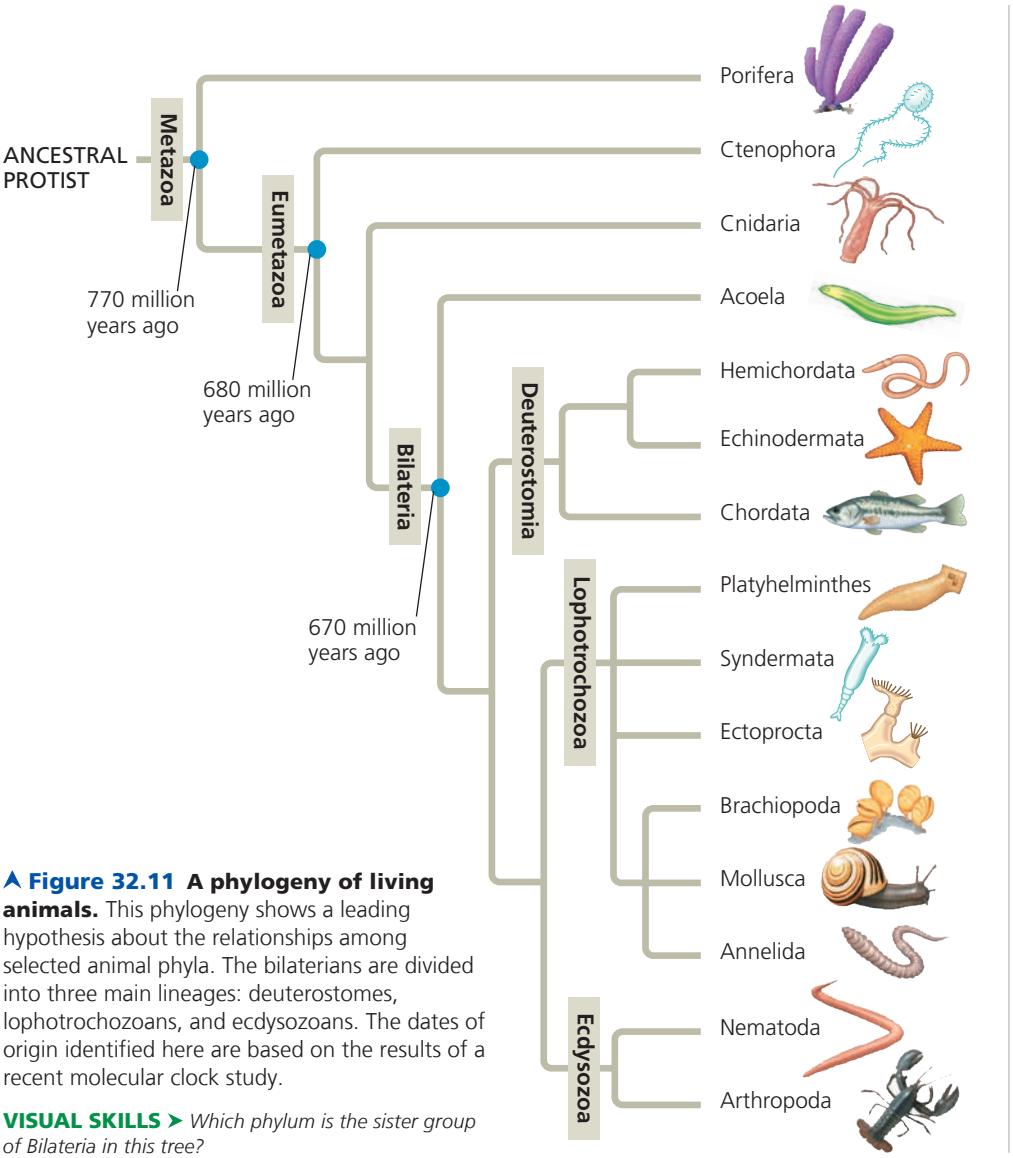
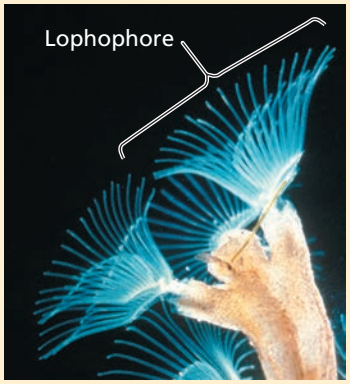
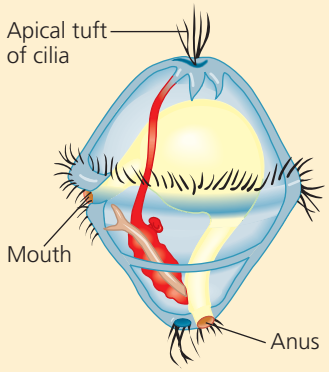


Figure 32.11 A phylogeny of living animals. This phylogeny shows a leading hypothesis about the relationships among selected animal phyla. The bilaterians are divided into three main lineages: deuterostomes, lophotrochozoans, and ecdysozoans. The dates of origin identified here are based on the results of a recent molecular clock study.

VISUAL SKILLS ➤ *Which phylum is the sister group of Bilateria in this tree?*



(a) Lophophore feeding structures of an ectoproct



(b) Structure of a trochophore larva

Figure 32.12 Morphological characteristics of lophotrochozoans.

cord; echinoderms lack these characteristics. These shared traits may have been present in the common ancestor of the deuterostome clade (and lost in the echinoderm lineage). As mentioned above, phylum Chordata, the only phylum with vertebrate members, also includes invertebrates.

Bilaterians also diversified in two major clades that are composed entirely of invertebrates: the *ecdysozoans* and the *lophotrochozoans*. The clade name **Ecdysozoa** refers to a characteristic shared by nematodes, arthropods, and some of the other ecdysozoan phyla that are not included in our survey. These animals secrete external skeletons (exoskeletons); the stiff covering of a cricket and the flexible cuticle of a nematode are examples. As the animal grows, it molts, squirming out of its old exoskeleton and secreting a larger one. The process of shedding the old exoskeleton is called *ecdysis*. Though named for this characteristic, the clade was proposed mainly on the basis of molecular data that support the common ancestry of its members. Furthermore, some taxa excluded from this clade by their molecular data, such as certain species of leeches, do in fact molt.

The name **Lophotrochozoa** refers to two different features observed in some animals belonging to this clade. Some lophotrochozoans, such as ectoprocts, develop a unique structure called a **lophophore** (from the Greek *lophos*, crest, and *pherein*, to carry), a crown of ciliated tentacles that function in feeding **(Figure 32.12a)**. Individuals in other phyla, including molluscs and annelids, go through a distinctive developmental stage called the **trochophore larva** **(Figure 32.12b)**—hence the name lophotrochozoan.

Future Directions in Animal Systematics

While many scientists think that current evidence supports the evolutionary relationships shown in Figure 32.11, aspects of this phylogeny continue to be debated. Although it can be frustrating that the phylogenies in textbooks cannot be

memorized as set-in-stone truths, the uncertainty inherent in these diagrams is a healthy reminder that science is an ongoing, dynamic process of inquiry. We'll conclude with three questions that are the focus of ongoing research.

1. **Are sponges monophyletic?** Traditionally, sponges were placed in a single phylum, Porifera. This view was challenged in the 1990s, when molecular studies indicated that sponges were paraphyletic; as a result, sponges were placed into several different phyla that branched near the base of the animal tree. Since 2009, however, several morphological and molecular studies have concluded that sponges are monophyletic after all, as traditionally thought and as shown in Figure 32.11. Researchers are currently sequencing the entire genomes of various sponges to investigate whether sponges are indeed monophyletic.
2. **Are ctenophores basal metazoans?** Many researchers have concluded that sponges are basal metazoans (see Figure 32.11). This conclusion was supported in a 2016 phylogenomic analysis, but several other recent studies have placed the comb jellies (phylum Ctenophora) at the base of the animal tree. In addition to the most recent phylogenomic results, data consistent with placing sponges at the base of the animal tree include fossil steroid evidence, molecular clock analyses, the morphological similarity of sponge collar cells to the cells of choanoflagellates (see Figure 32.3), and the fact that sponges are one of the few animal groups that lack tissues (as might be expected for basal animals). Ctenophores, on the other hand, have tissues, and their cells do not resemble the cells of choanoflagellates. At present, the idea that ctenophores are basal metazoans remains an intriguing but controversial hypothesis.
3. **Are acoelomate flatworms basal bilaterians?** A series of recent molecular papers have indicated that acoelomate flatworms (phylum Acoela) are basal bilaterians, as shown in Figure 32.11. A different conclusion was supported by a 2011 analysis, which placed acoelomates within Deuterostomia. Researchers are currently sequencing the genomes of several acoelomates and species from closely related groups to provide a more definitive test of the hypothesis that acoelomate flatworms are basal bilaterians. If further evidence supports this hypothesis, this would suggest that the bilaterians may have descended from a common ancestor that resembled living acoelomate flatworms—that is, from an ancestor that had a simple nervous system, a saclike gut with a single opening (the "mouth"), and no excretory system.

CONCEPT CHECK 32.4

1. Describe the evidence that cnidarians share a more recent common ancestor with other animals than with sponges.
2. **WHAT IF? ➤** Suppose ctenophores are basal metazoans and sponges are the sister group of all remaining animals. Under this hypothesis, redraw Figure 32.11 and discuss whether animals with tissues would form a clade.
3. **MAKE CONNECTIONS ➤** Based on the phylogeny in Figure 32.11 and the information in Figure 25.11, evaluate this statement: "The Cambrian explosion actually consists of three explosions, not one."

For suggested answers, see Appendix A.

32 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

CONCEPT 32.1

Animals are multicellular, heterotrophic eukaryotes with tissues that develop from embryonic layers *(pp. 726–727)*

VOCAB SELF-QUIZ goo.gl/Rn5Uax

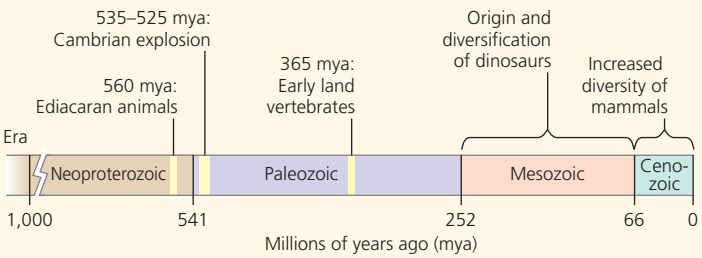
- Animals are heterotrophs that ingest their food.
- Animals are multicellular eukaryotes. Their cells are supported and connected to one another by collagen and other structural proteins located outside the cell membrane. Nervous tissue and muscle tissue are key animal features.
- In most animals, **gastrulation** follows the formation of the **blastula** and leads to the formation of embryonic tissue layers. Most animals have *Hox* genes that regulate the development of body form. Although *Hox* genes have been highly conserved over the course of evolution, they can produce a wide diversity of animal morphology.

? *Describe key ways that animals differ from plants and fungi.*

CONCEPT 32.2

The history of animals spans more than half a billion years *(pp. 727–731)*

- Fossil biochemical evidence and molecular clock analyses indicate that animals arose over 700 million years ago.
- Genomic analyses suggest that key steps in the origin of animals involved new ways of using proteins that were encoded by genes found in choanoflagellates.



? *What caused the Cambrian explosion? Describe current hypotheses.*

CONCEPT 32.3

Animals can be characterized by "body plans" *(pp. 731–734)*

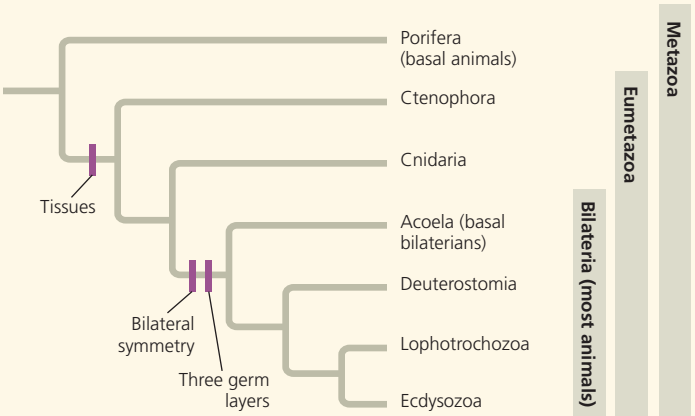
- Animals may lack symmetry or may have **radial** or **bilateral symmetry**. Bilaterally symmetrical animals have **dorsal** and **ventral** sides, as well as **anterior** and **posterior** ends.
- Eumetazoan embryos may be **diploblastic** (two germ layers) or **triploblastic** (three germ layers). Triploblastic animals with a **body cavity** may have a **pseudocoelom** or a true **coelom**.
- **Protostome** and **deuterostome development** often differ in patterns of cleavage, coelom formation, and **blastopore** fate.

? *Describe how body plans provide useful information yet should be interpreted cautiously as evidence of evolutionary relationships.*

CONCEPT 32.4

Views of animal phylogeny continue to be shaped by new molecular and morphological data *(pp. 734–736)*

- This phylogenetic tree shows key steps in animal evolution:



? *Consider clades Bilateria, Lophotrochozoa, Metazoa, Chordata, Ecdysozoa, Eumetazoa, and Deuterostomia. List the clades to which humans belong in order from the most to the least inclusive clade.*

TEST YOUR UNDERSTANDING

Multiple-choice Self-Quiz questions 1–4 can be found in the Study Area in MasteringBiology.

5. **EVOLUTION CONNECTION** A professor begins a lecture on animal phylogeny (as shown in Figure 32.11) by saying, "We are all worms." In this context, what did she mean?

6. **SCIENTIFIC INQUIRY • INTERPRET THE DATA** Redraw the bilaterian portion of Figure 32.11 for the nine phyla in the table below.

PRACTICE TEST
goo.gl/iAsVgL

Consider these blastopore fates: protostomy (mouth develops from the blastopore), deuterostomy (anus develops from the blastopore), or neither (the blastopore closes and the mouth develops elsewhere). Depending on the blastopore fate of its members, label each branch that leads to a phylum with P, D, N, or a combination of these letters. What is the ancestral blastopore fate? How many times has blastopore fate changed over the course of evolution? Explain.

Blastopore Fate	Phyla
Protostomy (P)	Platyhelminthes, Syndermata, Nematoda; most Mollusca, most Annelida; few Arthropoda
Deuterostomy (D)	Echinodermata, Chordata; most Arthropoda; few Mollusca, few Annelida
Neither (N)	Acoela

Data adapted from A. Hejnol and M. Martindale, The mouth, the anus, and the blastopore—open questions about questionable openings. In *Animal Evolution: Genomes, Fossils and Trees,* eds. D. T. J. Littlewood and M. J. Telford, Oxford University Press, pp. 33–40 (2009).

7. **WRITE ABOUT A THEME: INTERACTIONS** Animal life changed greatly during the Cambrian explosion, with some groups expanding in diversity and others declining. Write a short essay (100–150 words) interpreting these events as feedback regulation at the level of the biological community.

8. **SYNTHESIZE YOUR KNOWLEDGE**

This organism is an animal. What can you infer about its body structure and lifestyle (that might not be obvious from its appearance)? This animal has a protostome developmental pattern and a trochophore larva. Identify the major clades that this animal belongs to. Explain your selection, and describe when these clades originated and how they are related to one another.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

33 Invertebrates

Figure 33.1 What function do the thin blue "fingers" of this organism have?

KEY CONCEPTS

33.1 Sponges are basal animals that lack tissues

33.2 Cnidarians are an ancient phylum of eumetazoans

33.3 Lophotrochozoans, a clade identified by molecular data, have the widest range of animal body forms

33.4 Ecdysozoans are the most species-rich animal group

33.5 Echinoderms and chordates are deuterostomes

A Dragon Without a Backbone

Starting with its striking colors and surreal shape, the blue dragon (*Glaucus atlanticus*) in **Figure 33.1** is full of surprises. Its thin, finger-like structures increase the surface area of its body, aiding in respiration and helping it to float (upside down) on the sea's surface. This small sea slug also packs a powerful punch: It eats Portuguese men-of-war and absorbs their stinging cells, which the blue dragon then uses to deliver a deadly sting of its own.

Blue dragons are **invertebrates**: animals that lack a backbone. Invertebrates account for over 95% of known animal species. They occupy almost every habitat on Earth, from the scalding water released by deep-sea "black smoker" hydrothermal vents to the frozen ground of Antarctica. Evolution in these varied environments has produced an immense diversity of forms, including a species consisting of a flat bilayer of cells and species with features such as silk-spinning glands, pivoting spines, and tentacles covered with suction cups. Invertebrates range from microscopic organisms to organisms that can grow to 18 m long (1.5 times the length of a school bus).

In this chapter, we'll take a tour of the invertebrate world, using the phylogenetic tree in **Figure 33.2** as a guide. **Figure 33.3** surveys 23 invertebrate phyla as representatives of invertebrate diversity. Many of those phyla are explored in more detail in the rest of this chapter.

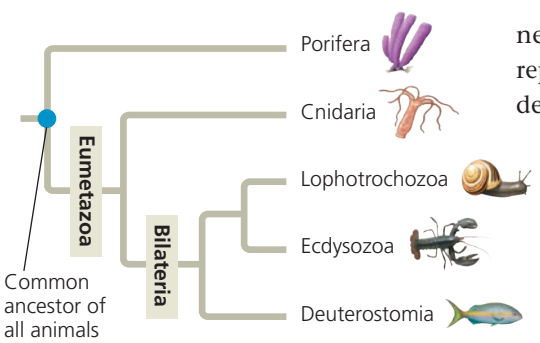


Figure 33.2 Review of animal phylogeny. Except for sponges (phylum Porifera) and a few other groups, all animals have tissues and are eumetazoans. Most animals are bilaterians (see Figure 32.11).

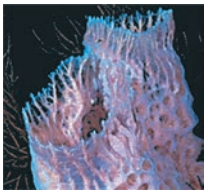
When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

Figure 33.3 Exploring Invertebrate Diversity

Kingdom Animalia encompasses 1.3 million known species, and estimates of the total number of species range as high as 10–20 million. Of the 23 phyla surveyed here, 12 are discussed more fully in this chapter, Chapter 32, or Chapter 34; cross-references are given at the end of their descriptions.

Porifera (5,500 species)



A sponge

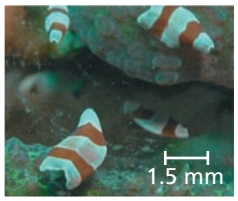
Animals in this phylum are informally called sponges. Sponges are sessile animals that lack tissues. They live as filter feeders, trapping particles that pass through the internal channels of their body (see Concept 33.1).

Cnidaria (10,000 species)

Cnidarians include corals, jellies, and hydras. These animals have a diploblastic, radially symmetrical body plan that includes a gastrovascular cavity with a single opening that serves as both mouth and anus (see Concept 33.2).

A jelly

Acoela (400 species)

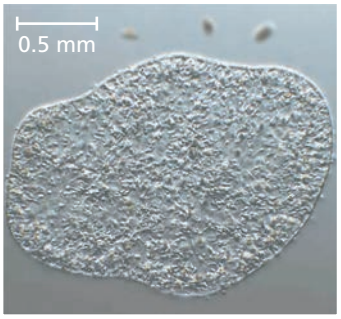


Acoel flatworms (LM)

Acoel flatworms have a simple nervous system and a saclike gut, and thus were once placed in phylum Platyhelminthes. Some molecular analyses, however, indicate that Acoela is a separate lineage that diverged before the three main bilaterian clades (see Concept 32.4).

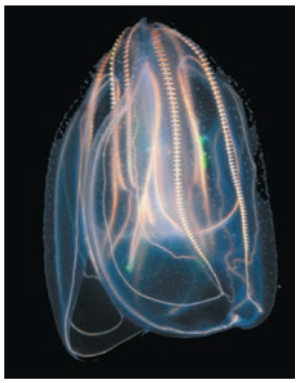
Placozoa (1 species)

The single known species in this phylum, *Trichoplax adhaerens*, doesn't even look like an animal. It consists of a simple bilayer of a few thousand cells. Placozoans are thought to be basal animals, but it is not yet known how they are related to Porifera, Cnidaria, and other phyla that diverged from most animals early in animal evolution. *Trichoplax* can reproduce by dividing into two individuals or by budding off many multicellular individuals.



A placozoan (LM)

Ctenophora (100 species)



A ctenophore, or comb jelly

Ctenophores (comb jellies) are diploblastic and radially symmetrical like cnidarians, suggesting that both phyla diverged from other animals very early (see Figure 32.11). Comb jellies make up much of the ocean's plankton. They have many distinctive traits, including eight "combs" of cilia that propel the animals through the water. When a small animal contacts the tentacles of some comb jellies, specialized cells burst open, covering the prey with sticky threads.

Lophotrochozoa

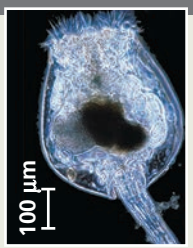
Platyhelminthes (20,000 species)

A marine flatworm

Flatworms (including tapeworms, planarians, and flukes) have bilateral symmetry and a central nervous system that processes information from sensory structures. They have no body cavity or specialized organs for circulation (see Concept 33.3).

Syndermata (2,900 species)

This recently established phylum includes two groups formerly classified as separate phyla: the rotifers, microscopic animals with complex organ systems, and the acanthocephalans, highly modified parasites of vertebrates (see Concept 33.3).



A rotifer (LM)

Ectoprocta (4,500 species)

Ectoprocts

Ectoprocts (also known as bryozoans) live as sessile colonies and are covered by a tough exoskeleton (see Concept 33.3).

Brachiopoda (335 species)

A brachiopod

Brachiopods, or lamp shells, may be easily mistaken for clams or other molluscs. However, most brachiopods have a unique stalk that anchors them to their substrate, as well as a crown of cilia called a lophophore (see Concept 33.3).

Continued on next page

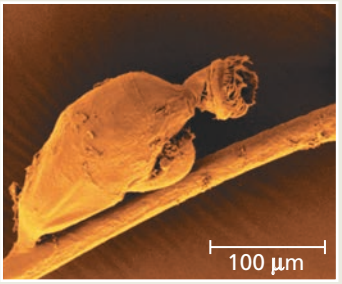
Lophotrochozoa (continued)

Gastrotricha (800 species)

Gastrotrichans are tiny worms whose ventral surface is covered with cilia, leading them to be called hairy bellies. Most species live on the bottoms of lakes or oceans, where they feed on small organisms and partially decayed organic matter. This individual has consumed algae, visible as the greenish material inside its gut.

A gastrotrichan (differential-interference-contrast LM)

Cycliophora (1 species)



A cycliophoran (colorized SEM)

The only known cycliphoran species, *Symbion pandora*, was discovered in 1995 on the mouthparts of a lobster. This tiny, vase-shaped creature has a unique body plan and a particularly bizarre life cycle. Males impregnate females that are still developing in their mothers' bodies. The fertilized females then escape, settle elsewhere on the lobster, and release their offspring. The offspring apparently leave that lobster and search for another one to which they attach.

A ribbon worm

Nemertea (900 species)

Also called proboscis worms or ribbon worms, nemerteans swim through water or burrow in sand, extending a unique proboscis to capture prey. Like flatworms, they lack a true coelom. However, unlike flatworms, nemerteans have an alimentary canal and a closed circulatory system in which the blood is contained in vessels and hence is distinct from fluid in the body cavity.

Annelida (16,500 species)

Annelids, or segmented worms, are distinguished from other worms by their body segmentation. Earthworms are the most familiar annelids, but the phylum consists primarily of marine and freshwater species (see Concept 33.3).

Mollusca (100,000 species)

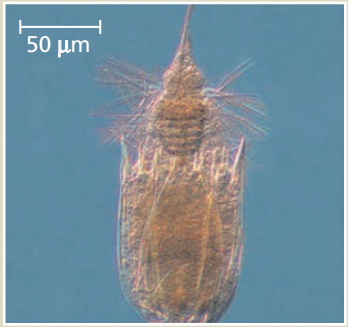
Molluscs (including snails, clams, squids, and octopuses) have a soft body that in many species is protected by a hard shell (see Concept 33.3).

An octopus

Ecdysozoa

Loricifera (10 species)

Loriciferans (from the Latin *lorica*, corset, and *ferre*, to bear) are tiny animals that inhabit sediments on the sea floor. A loriciferan can telescope its head, neck, and thorax in and out of the lorica, a pocket formed by six plates surrounding the abdomen. Though the natural history of loriciferans is mostly a mystery, at least some species likely eat bacteria.



A loriciferan (LM)

Priapula (16 species)

A priapulan

Priapulans are worms with a large, rounded proboscis at the anterior end. (They are named after Priapos, the Greek god of fertility, who was symbolized by a giant penis.) Ranging from 0.5 mm to 20 cm in length, most species burrow through seafloor sediments. Fossil evidence suggests that priapulans were among the major predators during the Cambrian period.

Ecdysozoa (continued)

Onychophora (110 species)

An onychophoran

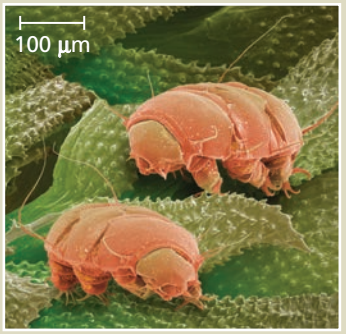
Onychophorans, also called velvet worms, originated during the Cambrian explosion (see Chapter 32). Originally, they thrived in the ocean, but at some point they succeeded in colonizing land. Today they live only in humid forests. Onychophorans have fleshy antennae and several dozen pairs of saclike legs.

Nematoda (25,000 species)

A roundworm

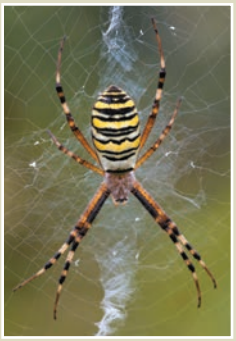
Also called roundworms, nematodes are enormously abundant and diverse in the soil and in aquatic habitats; many species parasitize plants and animals. Their most distinctive feature is a tough cuticle that coats their body (see Concept 33.4).

Tardigrada (800 species)



Tardigrades (colorized SEM)

Tardigrades (from the Latin *tardus*, slow, and *gradus*, step) are sometimes called water bears for their overall shape and lumbering, bearlike gait. Most tardigrades are less than 0.5 mm in length. Some live in oceans or fresh water, while others live on plants or animals. Harsh conditions may cause tardigrades to enter a state of dormancy; while dormant, they can survive for days at temperatures as low as –200°C! A 2015 phylogenomic study found that over 15% of tardigrade genes entered their genome by horizontal gene transfer, the largest fraction known for any animal.



A spider (an arachnid)

Arthropoda (1,000,000 species)

The vast majority of known animal species, including insects, crustaceans, and arachnids, are arthropods. All arthropods have a segmented exoskeleton and jointed appendages (see Concept 33.4).

Deuterostomia

Hemichordata (85 species)

An acorn worm

Like echinoderms and chordates, hemichordates are members of the deuterostome clade (see Chapter 32). Hemichordates share some traits with chordates, such as gill slits and a dorsal nerve cord. The largest group of hemichordates is the enteropneusts, or acorn worms. Acorn worms are marine and generally live buried in mud or under rocks; they may grow to more than 2 m in length.

Chordata (57,000 species)

More than 90% of all known chordate species have backbones (and thus are vertebrates). However, the phylum Chordata also includes two groups of invertebrates: lancelets and tunicates. See Chapter 34 for a full discussion of this phylum.

A tunicate

Echinodermata (7,000 species)

A sea urchin

Echinoderms, such as sand dollars, sea stars, and sea urchins, are marine animals in the deuterostome clade that are bilaterally symmetrical as larvae but not as adults. They move and feed by using a network of internal canals to pump water to different parts of their body (see Concept 33.5).

CONCEPT 33.1

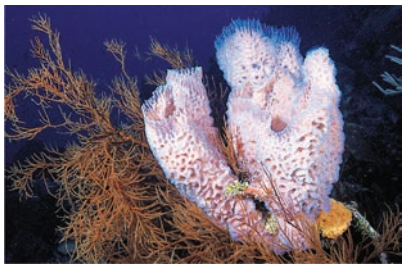
Sponges are basal animals that lack tissues

Animals in the phylum Porifera are known informally as sponges. (Recent molecular studies indicate that sponges are monophyletic, and that is the phylogeny we present here; this remains under debate, however, as some studies suggest that sponges are paraphyletic.) Among the simplest of animals, sponges are sedentary and were mistaken for plants by the ancient Greeks. Most species are marine, and they range in size from a few millimeters to a few meters. Sponges are **filter feeders**: They filter out food particles suspended in the surrounding water as they draw it through their body, which in some species resembles a sac perforated with pores. Water is drawn through the pores into a central cavity, the **spongocoel**, and then flows out of the sponge through a larger opening called the **osculum (Figure 33.4)**. More complex sponges have folded body walls, and many contain branched water canals and several oscula.

Sponges represent a lineage that diverged from other animals early in the history of the group; thus, they are said to be *basal animals*. Unlike nearly all other animals, sponges lack tissues, groups of similar cells that act as a functional unit as in muscle tissue and nervous tissue. However, the sponge body does contain several different cell types. For example, lining the interior of the spongocoel are flagellated **choanocytes**, or collar cells (named for the finger-like projections that form a "collar" around the flagellum). These cells engulf bacteria and other food particles by phagocytosis. The similarity between choanocytes and the cells of choanoflagellates supports molecular evidence suggesting that animals evolved from a choanoflagellate-like ancestor (see Figure 32.3).

The body of a sponge consists of two layers of cells separated by a gelatinous region called the **mesohyl**. Because both cell layers are in contact with water, processes such as gas exchange and waste removal can occur by diffusion across the membranes of these cells. Other tasks are performed by cells called **amoebocytes**, named for their use of pseudopodia. These cells move through the mesohyl and have many functions. For example, they take up food from the surrounding water and from choanocytes, digest it, and carry nutrients to other cells. Amoebocytes also manufacture tough skeletal fibers within the mesohyl. In some sponges,

Figure 33.4 Anatomy of a sponge. In the main diagram, portions of the front and back wall are cut away to show the sponge's internal structure.



Azure vase sponge (*Callyspongia plicifera*)

these fibers are sharp spicules made from calcium carbonate or silica. Other sponges produce more flexible fibers composed of a protein called spongin; you may have seen these pliant skeletons being sold as brown bath sponges. Finally, and perhaps most importantly, amoebocytes are *totipotent* (capable of becoming other types of sponge cells). This gives the sponge body remarkable flexibility, enabling it to adjust its shape in response to changes in its physical environment (such as the direction of water currents).

Most sponges are **hermaphrodites**, meaning that each individual functions as both male and female in sexual reproduction by producing sperm *and* eggs. Almost all sponges exhibit sequential hermaphroditism: They function first as one sex and then as the other. Cross-fertilization can result when sperm released into the water current by an individual functioning as a male is drawn into a neighboring individual that is functioning as a female. The resulting zygotes develop into flagellated, swimming larvae that disperse from the parent sponge. After settling on a suitable substrate, a larva develops into a sessile adult.

Sponges produce a variety of antibiotics and other defensive compounds, which hold promise for fighting human diseases. For example, a compound called cribrostatin isolated from marine sponges can kill both cancer cells and penicillin-resistant strains of the bacterium *Streptococcus*. Other sponge-derived compounds are also being tested as possible anticancer agents.

CONCEPT CHECK 33.1

1. Describe how sponges feed.
2. **WHAT IF?** ➤ Some molecular evidence suggests that the sister group of animals is not the choanoflagellates, but rather a group of parasitic protists, Mesomycetozoa. Given that these parasites lack collar cells, can this hypothesis be correct? Explain.

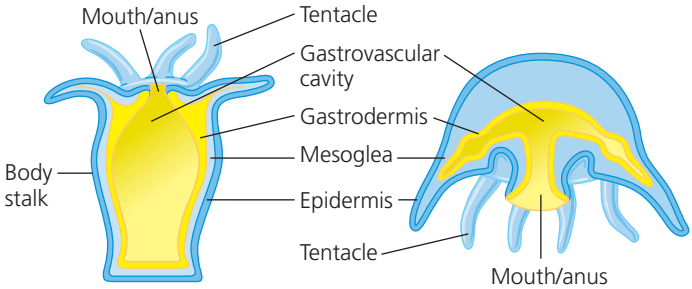
For suggested answers, see Appendix A.

CONCEPT 33.2

Cnidarians are an ancient phylum of eumetazoans

All animals except sponges and a few other groups are *eumetazoans* ("true animals"), members of a clade of animals with tissues. One of the first lineages to have diverged from others in this clade is the phylum Cnidaria, which originated about 680 million years ago according to DNA analyses. Cnidarians have diversified into a wide range of sessile and motile forms, including hydras, corals, and jellies (commonly called "jellyfish"). Yet most cnidarians still exhibit the relatively simple, diploblastic, radial body plan that existed in early members of the group some 560 million years ago.

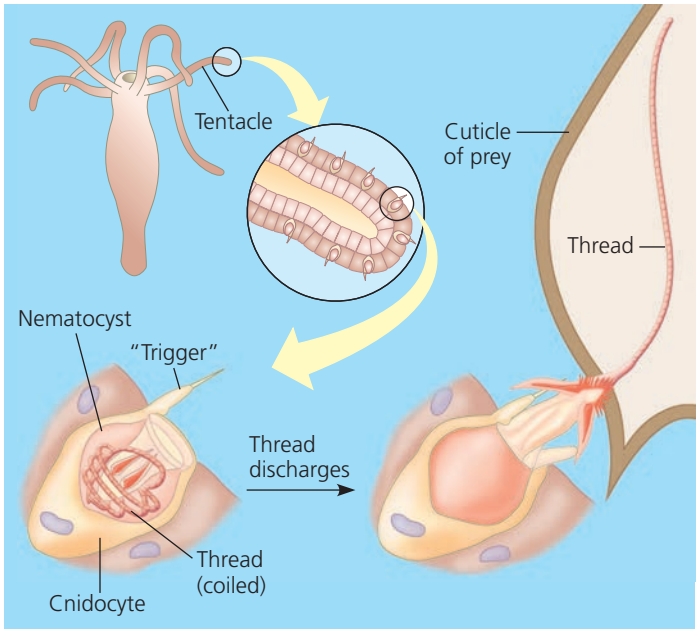
Figure 33.5 Polyp and medusa forms of cnidarians. The body wall of a cnidarian has two layers of cells: an outer layer of epidermis (darker blue; derived from ectoderm) and an inner layer of gastrodermis (yellow; derived from endoderm). Digestion begins in the gastrovascular cavity and is completed inside food vacuoles in the gastrodermal cells. Sandwiched between the epidermis and gastrodermis is a gelatinous layer, the mesoglea.



The basic body plan of a cnidarian is a sac with a central digestive compartment, the **gastrovascular cavity**. A single opening to this cavity functions as both mouth and anus. There are two variations on this body plan: the largely sessile polyp and the more motile medusa **(Figure 33.5)**. **Polyps** are cylindrical forms that adhere to the substrate by the aboral end of their body (the end opposite the mouth) and extend their tentacles, waiting for prey. Examples of the polyp form include hydras and sea anemones. Although they are primarily sedentary, many polyps can move slowly across their substrate using muscles at the aboral end of their body. When threatened by a predator, some sea anemones can detach from the substrate and "swim" by bending their body column back and forth, or thrashing their tentacles. A **medusa** (plural, *medusae*) resembles a flattened, mouth-down version of the polyp. It moves freely in the water by a combination of passive drifting and contractions of its bell-shaped body. Medusae include free-swimming jellies. The tentacles of a jelly dangle from the oral surface, which points downward. Some cnidarians exist only as polyps or only as medusae; others have both a polyp stage and a medusa stage in their life cycle.

Cnidarians are predators that often use tentacles arranged in a ring around their mouth to capture prey and push the food into their gastrovascular cavity, where digestion begins. Enzymes are secreted into the cavity, thus breaking down the prey into a nutrient-rich broth. Cells lining the cavity then absorb these nutrients and complete the digestive process; any undigested remains are expelled through the cnidarian's mouth/anus. The tentacles are armed with batteries of **cnidocytes**, cells unique to cnidarians that function in defense and prey capture. Cnidocytes contain *cnidae* (from the Greek *cnide*, nettle), capsule-like organelles that are capable of exploding outward and that give phylum Cnidaria its

Figure 33.6 A cnidocyte of a hydra. This type of cnidocyte contains a stinging capsule, the nematocyst, which contains a coiled thread. When a "trigger" is stimulated by touch or by certain chemicals, the thread shoots out, puncturing and injecting poison into prey.



name **(Figure 33.6)**. Specialized cnidae called **nematocysts** contain a stinging thread that can penetrate the body wall of the cnidarian's prey. Other kinds of cnidae have long threads that stick to or entangle small prey that bump into the cnidarian's tentacles.

Contractile tissues and nerves occur in their simplest forms in cnidarians. Cells of the epidermis (outer layer) and gastrodermis (inner layer) have bundles of microfilaments arranged into contractile fibers. The gastrovascular cavity acts as a hydrostatic skeleton (see Concept 50.6) against which the contractile cells can work. When a cnidarian closes its mouth, the volume of the cavity is fixed, and contraction of selected cells causes the animal to change shape. Cnidarians have no brain. Instead, movements are coordinated by a noncentralized *nerve net* that is associated with sensory structures distributed around the body. Thus, the animal can detect and respond to stimuli from all directions.

Fossil and molecular evidence suggests that early in its evolutionary history, the phylum Cnidaria diverged into two major clades, Medusozoa and Anthozoa **(Figure 33.7)**.

Medusozoans

All cnidarians that produce a medusa are members of clade Medusozoa, a group that includes the *scyphozoans* (jellies) and *cubozoans* (box jellies) shown in Figure 33.7a, along with the *hydrozoans*. Most hydrozoans alternate between the polyp and medusa forms, as seen in the life cycle of *Obelia* **(Figure 33.8)**. The polyp stage, a colony of interconnected polyps in the case of *Obelia*, is more conspicuous than the

Figure 33.7 Cnidarians.

(a) Medusozoans

Many jellies are bioluminescent. Food captured by nematocyst-bearing tentacles is transferred to specialized oral arms (that lack nematocysts) for transport to the mouth.

This sea wasp produces a poison that can subdue fish, crustaceans (as seen here), and other large prey. The poison is more potent than cobra venom.

(b) Anthozoans

Sea anemones and other anthozoans exist only as polyps. Many anthozoans form symbiotic relationships with photosynthetic algae.

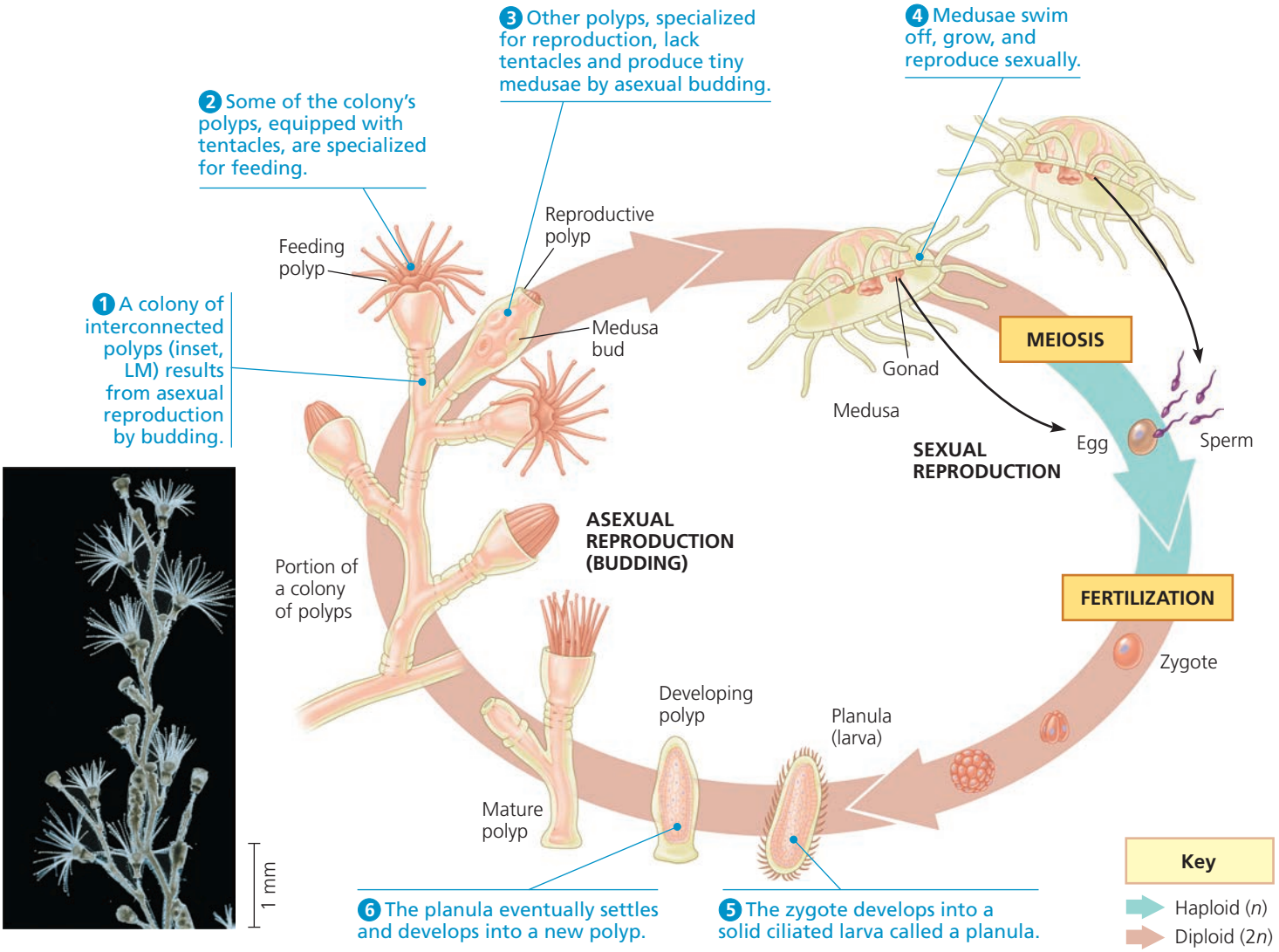
These star corals live as colonies of polyps. Their soft bodies are enclosed at the base by a hard exoskeleton.

Video: Jelly Swimming
Video: *Hydra* eating *Daphnia*

medusa. Hydras, among the few cnidarians found in fresh water, are also unusual hydrozoans in that they exist only in polyp form.

Unlike hydrozoans, most scyphozoans and cubozoans spend the majority of their life cycles in the medusa stage. Coastal scyphozoans, for example, often have a brief polyp stage during their life cycle, whereas those that live in the open ocean generally lack the polyp stage altogether. As their name (which means "cube animals") suggests, cubozoans have a box-shaped medusa stage. Most cubozoans live in tropical oceans and are equipped with highly toxic cnidocytes. For example, the sea wasp (*Chironex fleckeri*), a cubozoan that lives off the coast of northern Australia, is one of the deadliest organisms known: Its sting causes intense pain

Figure 33.8 The life cycle of the hydrozoan *Obelia*. The polyp is asexual, and the medusa is sexual, releasing eggs and sperm. These two stages alternate, one producing the other.



MAKE CONNECTIONS ➤ *Compare and contrast the* Obelia *life cycle to the life cycles in Figure 13.6. Which life cycle in that figure is most similar to that of* Obelia*? Explain. (See also Figure 29.3.)*

and can lead to respiratory failure, cardiac arrest, and death within minutes.

Anthozoans

Sea anemones and corals belong to the clade Anthozoa (see Figure 33.7b). These cnidarians occur only as polyps. Corals live as solitary or colonial forms, often forming symbioses with algae. Many species secrete a hard **exoskeleton** (external skeleton) of calcium carbonate. Each polyp generation builds on the skeletal remains of earlier generations, constructing rocklike reefs with shapes characteristic of their species. These skeletons are what we usually think of as coral.

Coral reefs are to tropical seas what rain forests are to tropical land areas: They provide habitat for many other species. Unfortunately, these reefs are being destroyed at an alarming rate. Pollution, overharvesting, and ocean acidification (see Figure 3.12) are major threats; global warming is likely also contributing to their demise by raising seawater temperatures above the range in which corals thrive.

CONCEPT CHECK 33.2

1. Compare and contrast the polyp and medusa forms of cnidarians.
2. **VISUAL SKILLS ➤** Use the cnidarian life cycle diagram in Figure 33.8 to determine the ploidy of a feeding polyp and of a medusa.
3. **MAKE CONNECTIONS ➤** Many new animal body plans emerged during and after the Cambrian explosion. In contrast, cnidarians today retain the same diploblastic, radial body plan found in cnidarians 560 million years ago. Are cnidarians therefore less successful or less "highly evolved" than other animal groups? Explain. (See Concepts 25.3 and 25.6.)

For suggested answers, see Appendix A.

CONCEPT 33.3

Lophotrochozoans, a clade identified by molecular data, have the widest range of animal body forms

The vast majority of animal species belong to the clade Bilateria, whose members exhibit bilateral symmetry and triploblastic development (see Concept 32.3). Most bilaterians also have a digestive tract with two openings (a mouth and an anus) and a coelom. Recent DNA analyses suggest that the common ancestor of living bilaterians lived about 670 million years ago. To date, however, the oldest fossil that is widely accepted as a bilaterian is of *Kimberella*, a mollusc (or close relative) that lived 560 million years ago (see Figure 32.5). Many other bilaterian groups first appeared in the fossil record during the Cambrian explosion (535 to 525 million years ago).

Molecular evidence suggests that today there are three major clades of bilaterally symmetrical animals: Lophotrochozoa, Ecdysozoa, and Deuterostomia. This section will focus on the first of these clades, the lophotrochozoans. Concepts 33.4 and 33.5 will explore the other two clades.

Although the clade Lophotrochozoa was identified by molecular data, its name comes from features found in some of its members. Some lophotrochozoans develop a structure called a *lophophore*, a crown of ciliated tentacles that functions in feeding, while others go through a distinctive stage called the *trochophore larva* (see Figure 32.12). Other members of the group have neither of these features. Few other unique morphological features are widely shared within the group—in fact, the lophotrochozoans are the most diverse bilaterian clade in terms of body plan. This diversity in form is reflected in the number of phyla classified in the group: Lophotrochozoa includes 18 phyla, more than twice the number in any other clade of bilaterians.

We'll now introduce six of the diverse lophotrochozoan phyla: the flatworms, rotifers and acanthocephalans, ectoprocts, brachiopods, molluscs, and annelids.

Flatworms

Flatworms (phylum Platyhelminthes) live in marine, freshwater, and damp terrestrial habitats. In addition to free-living species, flatworms include many parasitic species, such as flukes and tapeworms. Flatworms are so named because they have thin bodies that are flattened dorsoventrally (between the dorsal and ventral surfaces); the word *platyhelminth* means "flat worm." (Note that *worm* is not a formal taxonomic name but rather refers to a grade of animals with long, thin bodies.) The smallest flatworms are nearly microscopic free-living species, while some tapeworms are more than 20 m long.

Although flatworms undergo triploblastic development, they are *acoelomates* (animals that lack a body cavity). Their flat shape increases their surface area, placing all their cells close to water in the surrounding environment or in their gut. Because of this proximity to water, gas exchange and the elimination of nitrogenous waste (ammonia) can occur by diffusion across the body surface. As shown in **Figure 33.9**, a flat shape is one of several structural features that maximize surface area and have arisen (by convergent evolution) in different groups of animals and other organisms.

As you might expect since all their cells are close to water, flatworms have no organs specialized for gas exchange, and their relatively simple excretory apparatus functions mainly to maintain osmotic balance with their surroundings. This apparatus consists of **protonephridia**, networks of tubules with ciliated structures called *flame bulbs* that pull fluid through branched ducts opening to the outside (see Figure 44.9). Most flatworms have a gastrovascular cavity with only one opening. Though flatworms lack a circulatory system, the fine branches of the gastrovascular cavity distribute food directly to the animal's cells.

Early in their evolutionary history, flatworms separated into two lineages, Catenulida and Rhabditophora. Catenulida is a small clade of about 100 flatworm species, most of which live in freshwater habitats. Catenulids typically reproduce asexually by budding at their posterior end. The offspring often produce their own buds before detaching from the parent, thereby forming a chain of two to four genetically identical individuals—hence their informal name, "chain worms."

The other ancient flatworm lineage, Rhabditophora, is a diverse clade of about 20,000 freshwater and marine species, one example of which is shown in Figure 33.9. We'll explore the rhabditophorans in more detail, focusing on free-living and parasitic members of this clade.

Free-Living Species

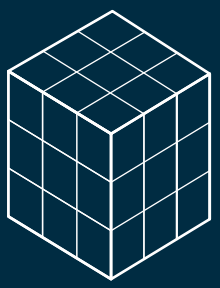
Free-living rhabditophorans are important as predators and scavengers in a wide range of freshwater and marine habitats. The best-known members of this group are freshwater species in the genus *Dugesia*, commonly called **planarians**. Abundant in unpolluted ponds and streams, planarians prey on smaller animals or feed on dead animals. They move by using cilia on their ventral surface, gliding along a film of mucus they secrete. Some other rhabditophorans also use their muscles to swim through water with an undulating motion.

A planarian's head features a pair of light-sensitive eyespots as well as lateral flaps that function mainly to detect specific chemicals. The planarian nervous system is more complex and centralized than the nerve nets of cnidarians

Figure 33.9 **MAKE CONNECTIONS**

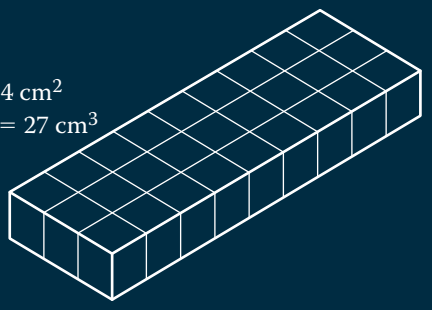
Maximizing Surface Area

In general, the amount of metabolic or chemical activity an organism can carry out is proportional to its mass or volume. Maximizing metabolic rate, however, requires the efficient uptake of energy and raw materials, such as nutrients and oxygen, as well as the effective disposal of waste products. For large cells, plants, and animals, these exchange processes have the potential to be limiting due to simple geometry. When a cell or organism grows without changing shape, its volume increases more rapidly than its surface area (see Figure 7.7). As a result, there is proportionately less surface area over which exchange processes can occur. The challenge posed by the relationship of surface area and volume occurs in diverse contexts and organisms, but the evolutionary adaptations that meet this challenge are similar. Structures that maximize surface area through flattening, folding, branching, and projections have an essential role in biological systems.



These diagrams compare surface area (SA) for two different shapes with the same volume (V). Note which shape has the greater surface area.

SA: $6\ (3\ \text{cm} \times 3\ \text{cm}) = 54\ \text{cm}^2$
V: $3\ \text{cm} \times 3\ \text{cm} \times 3\ \text{cm} = 27\ \text{cm}^3$



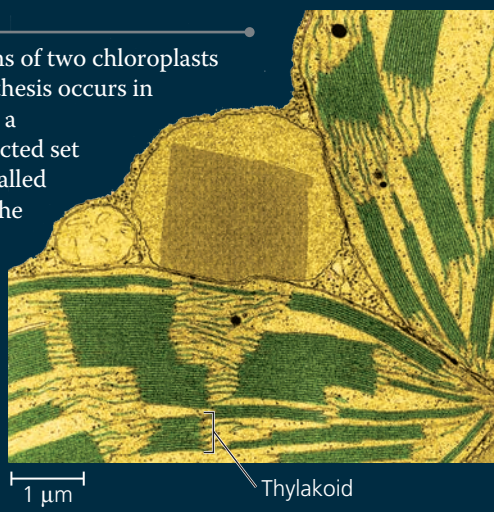
SA: $2\ (3\ \text{cm} \times 1\ \text{cm}) + 2\ (9\ \text{cm} \times 1\ \text{cm}) + 2\ (3\ \text{cm} \times 9\ \text{cm}) = 78\ \text{cm}^2$
V: $1\ \text{cm} \times 3\ \text{cm} \times 9\ \text{cm} = 27\ \text{cm}^3$

Flattening

By having a body that is only a few cells thick, an organism such as this flatworm can use its entire body surface for exchange. (See Figure 40.3.)

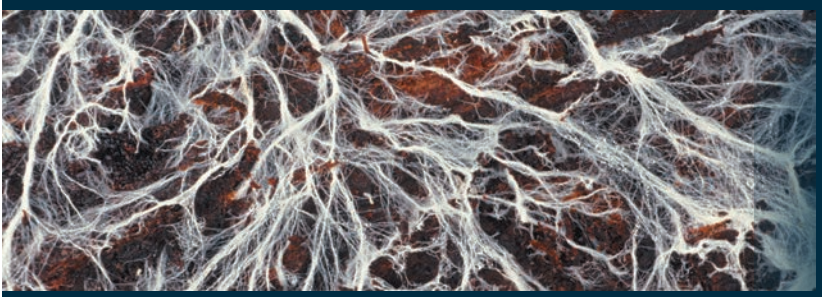
Folding

This TEM shows portions of two chloroplasts in a plant leaf. Photosynthesis occurs in chloroplasts, which have a flattened and interconnected set of internal membranes called thylakoid membranes. The foldings of the thylakoid membranes increase their surface area, enhancing the exposure to light and thus increasing the rate of photosynthesis. (See Figure 11.4.)



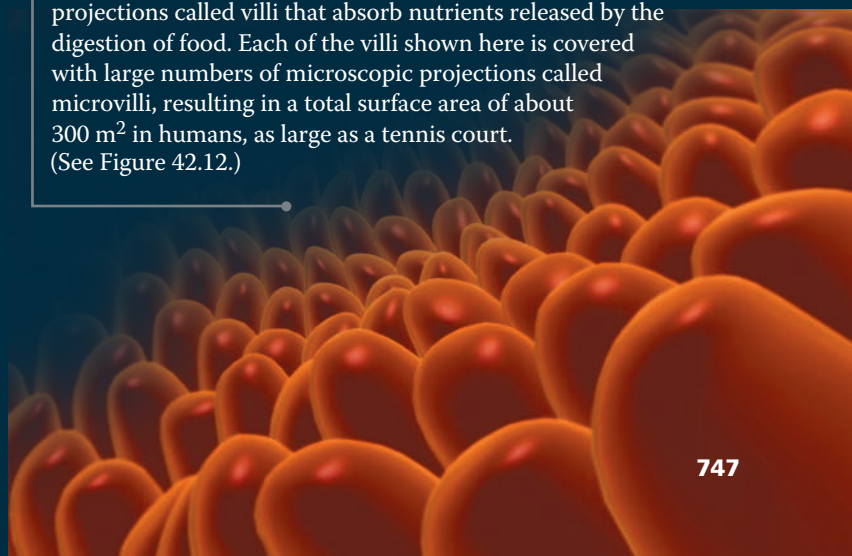
Branching

Water uptake relies on passive diffusion. The highly branched filaments of a fungal mycelium increase the surface area across which water and minerals can be absorbed from the environment. (See Figure 31.2.)



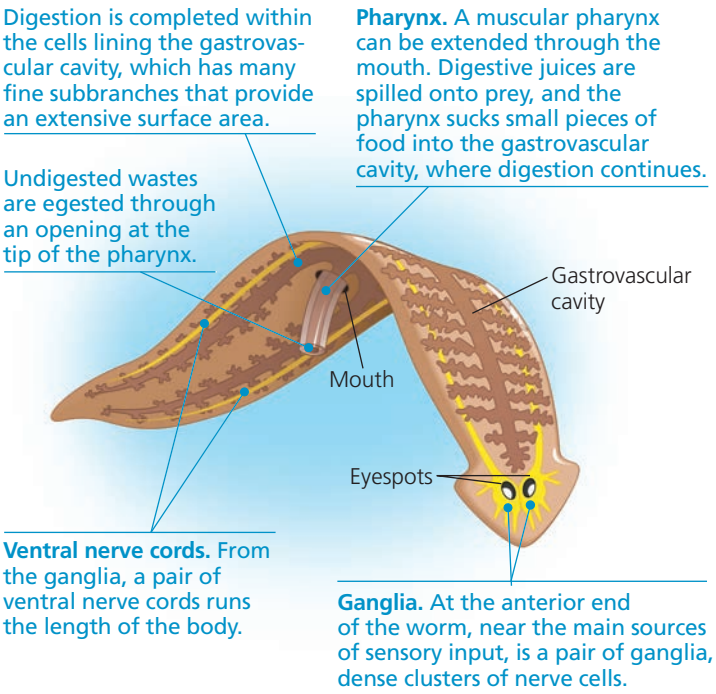
Projections

In vertebrates, the small intestine is lined with finger-like projections called villi that absorb nutrients released by the digestion of food. Each of the villi shown here is covered with large numbers of microscopic projections called microvilli, resulting in a total surface area of about 300 m^2 in humans, as large as a tennis court. (See Figure 42.12.)



MAKE CONNECTIONS ➤ *Find other examples of flattening, folding, branching, and projections (see Chapters 7, 10, 35, and 43). How is maximizing surface area important to the structure's function in each example?*

Figure 33.10 Anatomy of a planarian.



(Figure 33.10). Experiments have shown that planarians can learn to modify their responses to stimuli.

Some planarians can reproduce asexually through fission. The parent constricts roughly in the middle of its body, separating into a head end and a tail end; each end then regenerates the missing parts. Sexual reproduction also occurs. Planarians are hermaphrodites, and copulating mates typically cross-fertilize each other.

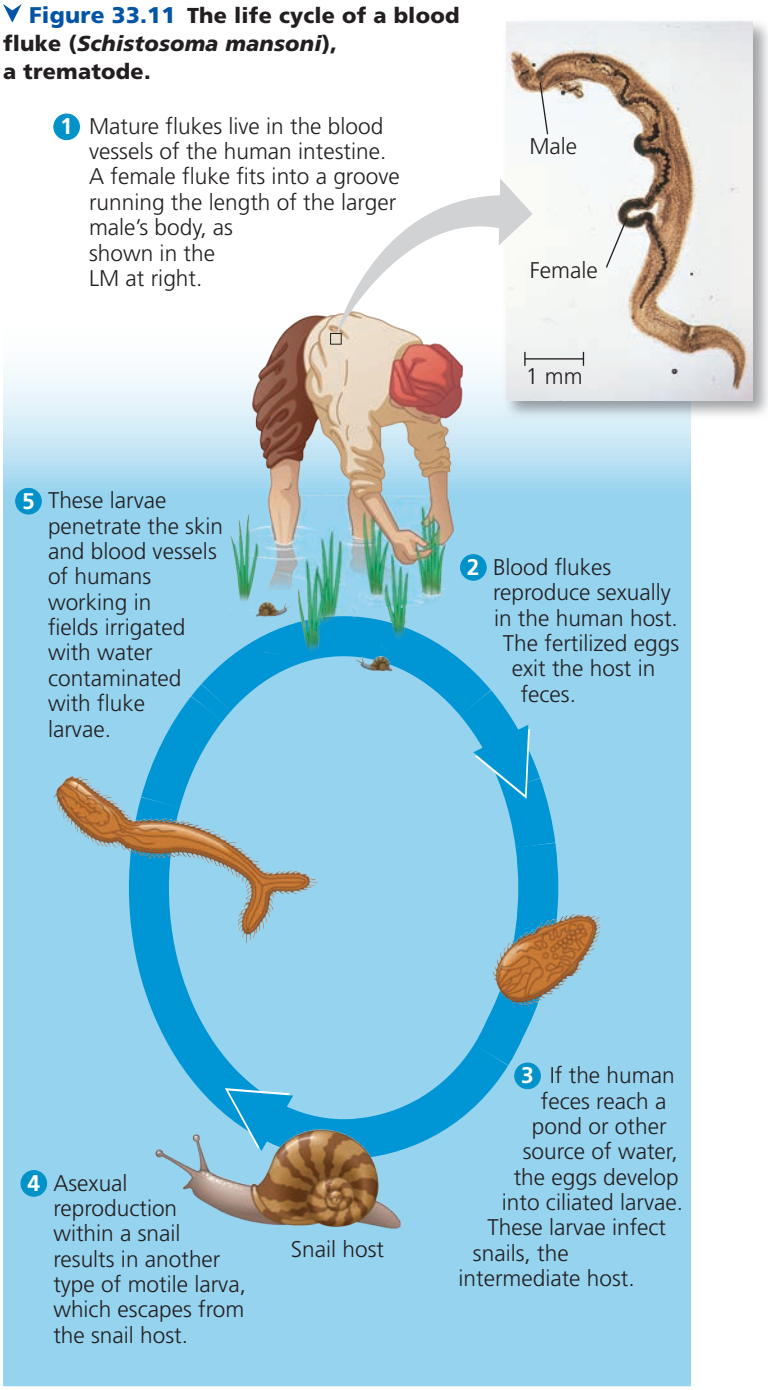
Parasitic Species

More than half of the known species of rhabditophorans live as parasites in or on other animals. Many have suckers that attach to the internal organs or outer surfaces of the host animal. In most species, a tough covering helps protect the parasites within their hosts. We'll discuss two ecologically and economically important subgroups of parasitic rhabditophorans, the trematodes and the tapeworms.

Trematodes As a group, trematodes parasitize a wide range of hosts, and most species have complex life cycles with alternating sexual and asexual stages. Many trematodes require an intermediate host in which larvae develop before infecting the final host (usually a vertebrate), where the adult worms live. For example, various trematodes that parasitize humans spend part of their lives in snail hosts **(Figure 33.11)**. Around the world, about 200 million people are infected with trematodes called blood flukes (*Schistosoma*) and suffer from schistosomiasis, a disease whose symptoms include pain, anemia, and diarrhea.

Living within more than one kind of host puts demands on trematodes that free-living animals don't face. A blood fluke,

Figure 33.11 The life cycle of a blood fluke (*Schistosoma mansoni*), a trematode.



WHAT IF? ➤ *Snails eat algae, whose growth is stimulated by nutrients found in fertilizer. How would the contamination of irrigation water with fertilizer likely affect the occurrence of schistosomiasis? Explain.*

for instance, must evade the immune systems of both snails and humans. By mimicking the surface proteins of its hosts, the blood fluke creates a partial immunological camouflage for itself. It also releases molecules that manipulate the hosts' immune systems into tolerating the parasite's existence. These defenses are so effective that individual blood flukes can survive in humans for more than 40 years.

Tapeworms The tapeworms are a second large and diverse group of parasitic rhabditophorans **(Figure 33.12)**. The adults

live mostly inside vertebrates, including humans. In many tapeworms, the anterior end, or scolex, is armed with suckers and often hooks that the worm uses to attach itself to the intestinal lining of its host. Tapeworms lack a mouth and gastrovascular cavity; they simply absorb nutrients released by digestion in the host's intestine. Absorption occurs across the tapeworm's body surface.

Posterior to the scolex is a long ribbon of units called proglottids, which are little more than sacs of sex organs. After sexual reproduction, proglottids loaded with thousands of fertilized eggs are released from the posterior end of a tapeworm and leave the host's body in feces. In one type of life cycle, feces carrying the eggs contaminate the food or water of intermediate hosts, such as pigs or cattle, and the tapeworm eggs develop into larvae that encyst in muscles of these animals. A human acquires the larvae by eating undercooked meat containing the cysts, and the worms develop into mature adults within the human. Large tapeworms can block the intestines and rob enough nutrients from the human host to cause nutritional deficiencies. Several different oral medications can kill the adult worms.

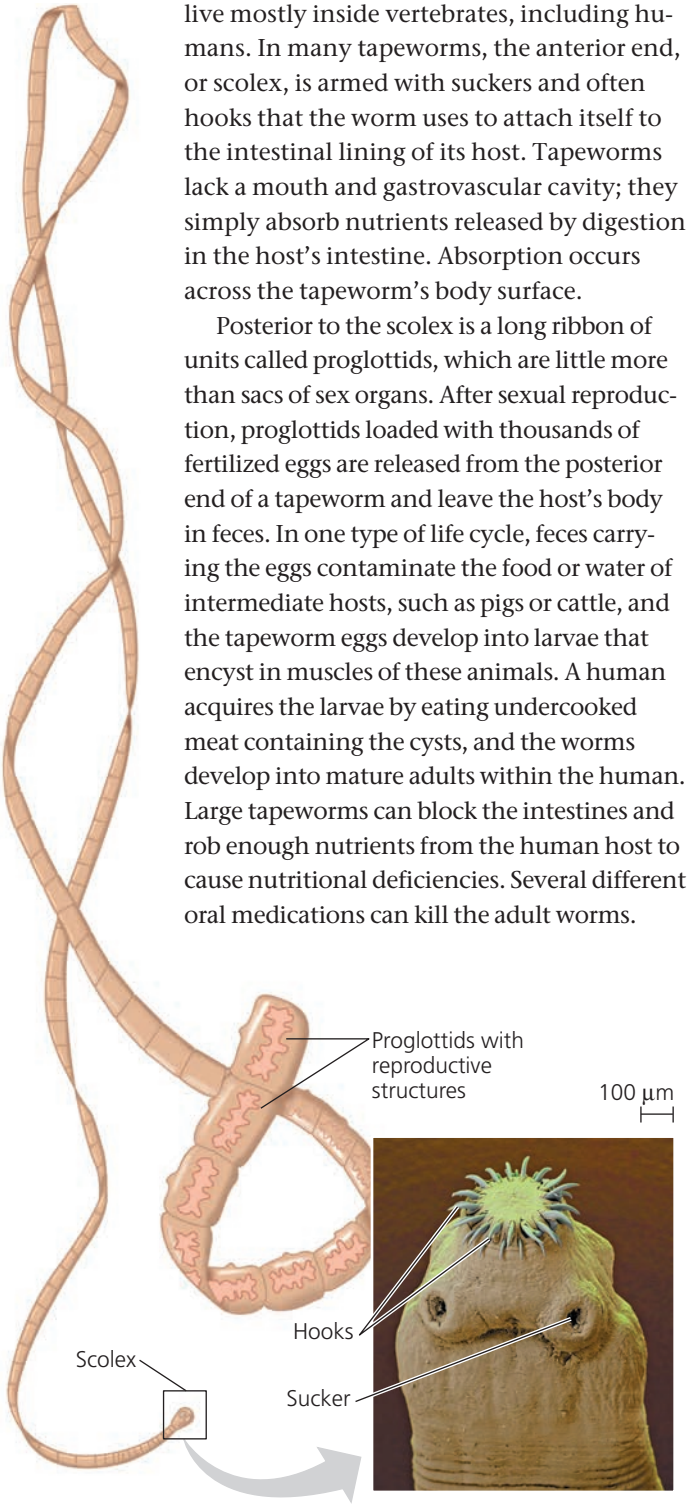
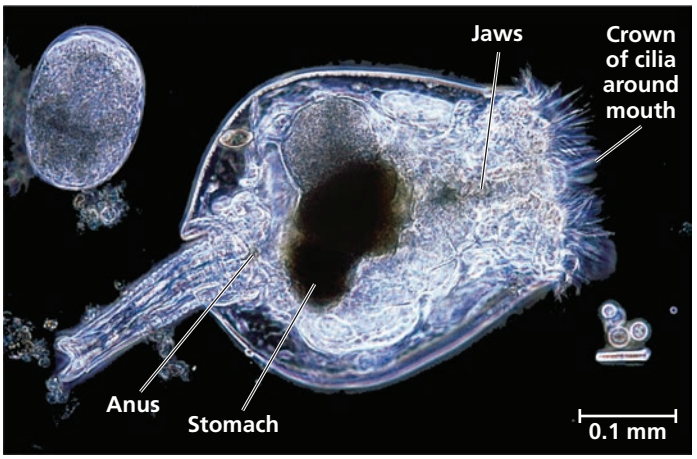


Figure 33.12 Anatomy of a tapeworm. The inset shows a close-up of the scolex (colorized SEM).

Rotifers and Acanthocephalans

Recent phylogenetic analyses have shown that two traditional animal phyla, the rotifers (former phylum Rotifera) and the acanthocephalans (former phylum Acanthocephala), should be combined into a single phylum, Syndermata. Each of the two groups has distinctive characteristics.

Figure 33.13 A rotifer. These pseudocoelomates, smaller than many protists, are generally more anatomically complex than flatworms (LM).



Rotifers

There are roughly 1,800 species of rotifers, tiny animals that inhabit freshwater, marine, and damp soil habitats. Ranging in size from about 50 µm to 2 mm, rotifers are smaller than many protists but nevertheless are multicellular and have specialized organ systems **(Figure 33.13)**. In contrast to cnidarians and flatworms, which have a gastrovascular cavity, rotifers have an **alimentary canal**, a digestive tube with two openings, a mouth and an anus. Internal organs lie within the *pseudocoelom*, a body cavity that is not completely lined by mesoderm (see Figure 32.9b). Fluid in the pseudocoelom serves as a hydrostatic skeleton. Movement of a rotifer's body distributes the fluid throughout the body, circulating nutrients.

The word *rotifer* is derived from the Latin meaning "wheel-bearer," a reference to the crown of cilia that draws a vortex of water into the mouth. Posterior to the mouth, rotifers have jaws called trophi that grind up food, mostly microorganisms suspended in the water. Digestion is then completed farther along the alimentary canal. Most other bilaterians also have an alimentary canal, which enables the stepwise digestion of a wide range of food particles.

Rotifers exhibit some unusual forms of reproduction. Some species consist only of females that produce more females from unfertilized eggs, a type of asexual reproduction called **parthenogenesis**. Some other invertebrates (for example, aphids and some bees) and even some vertebrates (for example, some lizards and some fishes) can also reproduce in this way. In addition to being able to produce females by parthenogenesis, some rotifers can also reproduce sexually under certain conditions, such as high levels of crowding. The resulting embryos can remain dormant for years. Once they break dormancy, the embryos develop into another generation of females that reproduce asexually.

It is puzzling that many rotifer species persist without males. The vast majority of animals and plants reproduce sexually at least some of the time, and sexual reproduction has

certain advantages over asexual reproduction (see Concept 45.1). For example, species that reproduce asexually tend to accumulate harmful mutations in their genomes faster than sexually reproducing species. As a result, asexual species should experience higher rates of extinction.

Seeking to understand how they persist without males, researchers have been studying a clade of asexual rotifers named Bdelloidea. Some 360 species of bdelloid rotifers are known, and all of them reproduce by parthenogenesis without any males. Paleontologists have discovered bdelloid rotifers preserved in 35-million-year-old amber, and the morphology of these fossils resembles only the female form, with no evidence of males. Molecular clock analyses indicate that bdelloids have been asexual for over 50 million years. While it appears that they do not reproduce sexually, bdelloid rotifers may be able to generate genetic diversity in other ways. For example, bdelloids can tolerate very high levels of desiccation. When conditions improve and their cells rehydrate, DNA from other species enters their cells through cracks in the plasma membrane. Recent evidence suggests that this foreign DNA can be incorporated into the bdelloids' genome, thereby leading to increased genetic diversity.

Acanthocephalans

Acanthocephalans (1,100 species) are sexually reproducing parasites of vertebrates that lack a complete digestive tract and usually are less than 20 cm long. They are commonly called spiny-headed worms because of the curved hooks on the proboscis at the anterior end of their body **(Figure 33.14)**. Although they once were placed in their own phylum, recent studies have shown that acanthocephalans originated from within the group traditionally known as Rotifera. In particular, rotifers in the genus *Seison* share a more recent common ancestor with acanthocephalans than they do with other rotifers, making the acanthocephalans a group of highly modified "rotifers."

All acanthocephalans are parasites that have complex life cycles with two or more hosts. Some species manipulate the behavior of their intermediate hosts (generally arthropods) in ways that increase their chances of reaching their final hosts (generally vertebrates). For example, acanthocephalans that infect New Zealand mud crabs cause their hosts to move to more visible areas on the beach, where the crabs are more likely to be eaten by birds, the worms' final hosts.

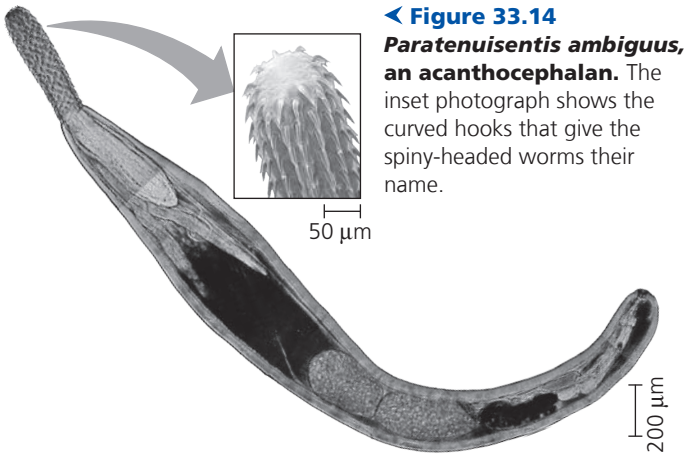


Figure 33.14 ***Paratenuisentis ambiguus*, an acanthocephalan.** The inset photograph shows the curved hooks that give the spiny-headed worms their name.

Lophophorates: Ectoprocts and Brachiopods

Bilaterians in the phyla Ectoprocta and Brachiopoda are among those known as lophophorates. These animals have a *lophophore*, a crown of ciliated tentacles around their mouth (see Figure 32.12a). As the cilia draw water toward the mouth, the tentacles trap suspended food particles. Other similarities, such as a U-shaped alimentary canal and the absence of a distinct head, reflect these organisms' sessile existence. In contrast to flatworms, which lack a body cavity, and rotifers, which have a pseudocoelom, lophophorates are *coelomates*, organisms with a body cavity that is completely lined by mesoderm (see Figure 32.9a).

Ectoprocts (from the Greek *ecto*, outside, and *procta*, anus) are colonial animals that superficially resemble clumps of moss. (In fact, their common name, bryozoans, means "moss animals.") In most species, the colony is encased in a hard exoskeleton studded with pores through which the lophophores extend **(Figure 33.15a)**. Most ectoproct species live in the sea, where they are among the most widespread and numerous sessile animals. Several species are important reef builders. Ectoprocts also live in lakes and rivers. Colonies of the freshwater ectoproct *Pectinatella magnifica* grow on submerged sticks or rocks and can grow into a gelatinous, ball-shaped mass more than 10 cm across.

Brachiopods, or lamp shells, superficially resemble clams and other hinge-shelled molluscs, but the two halves of the brachiopod shell are dorsal and ventral rather than lateral, as in clams **(Figure 33.15b)**. All brachiopods are marine. Most live attached to the seafloor by a stalk, opening their shell slightly to allow water to flow through the lophophore. The living brachiopods are remnants of a much richer past that included 30,000 species in the Paleozoic and Mesozoic eras. Some living brachiopods, such as those in the

Figure 33.15 Lophophorates.

(a) Ectoprocts, such as this creeping bryozoan (*Plumatella repens*), are colonial lophophorates.

(b) Brachiopods, such as this lampshell (*Terebratulina retusa*), have a hinged shell. The two parts of the shell are dorsal and ventral.

genus *Lingula*, appear nearly identical to fossils of species that lived 400 million years ago.

Molluscs

Snails and slugs, oysters and clams, and octopuses and squids are all molluscs (phylum Mollusca). There are over 100,000 known species, making them the second most diverse phylum of animals (after the arthropods, discussed later). Although the majority of molluscs are marine, roughly 8,000 species inhabit fresh water, and 28,000 species of snails and slugs live on land. All molluscs are soft-bodied, and most secrete a hard protective shell made of calcium carbonate. Slugs, squids, and octopuses have a reduced internal shell or have lost their shell completely during their evolution.

Despite their apparent differences, all molluscs have a similar body plan **(Figure 33.16)**. Molluscs are coelomates, and their bodies have three main parts: a muscular **foot**, usually used for movement; a **visceral mass** containing most of the internal organs; and a **mantle**, a fold of tissue that drapes over the visceral mass and secretes a shell (if one is present). In many molluscs, the mantle extends beyond the visceral mass, producing a water-filled chamber, the **mantle cavity**, which houses the gills, anus, and excretory pores. Many molluscs feed by using a straplike organ called a **radula** to scrape up food.

Most molluscs have separate sexes, and their gonads (ovaries or testes) are located in the visceral mass. Many snails, however, are hermaphrodites. The life cycle of many marine molluscs includes a ciliated larval stage, the trochophore (see Figure 32.12b), which is also characteristic of marine annelids (segmented worms) and some other lophotrochozoans.

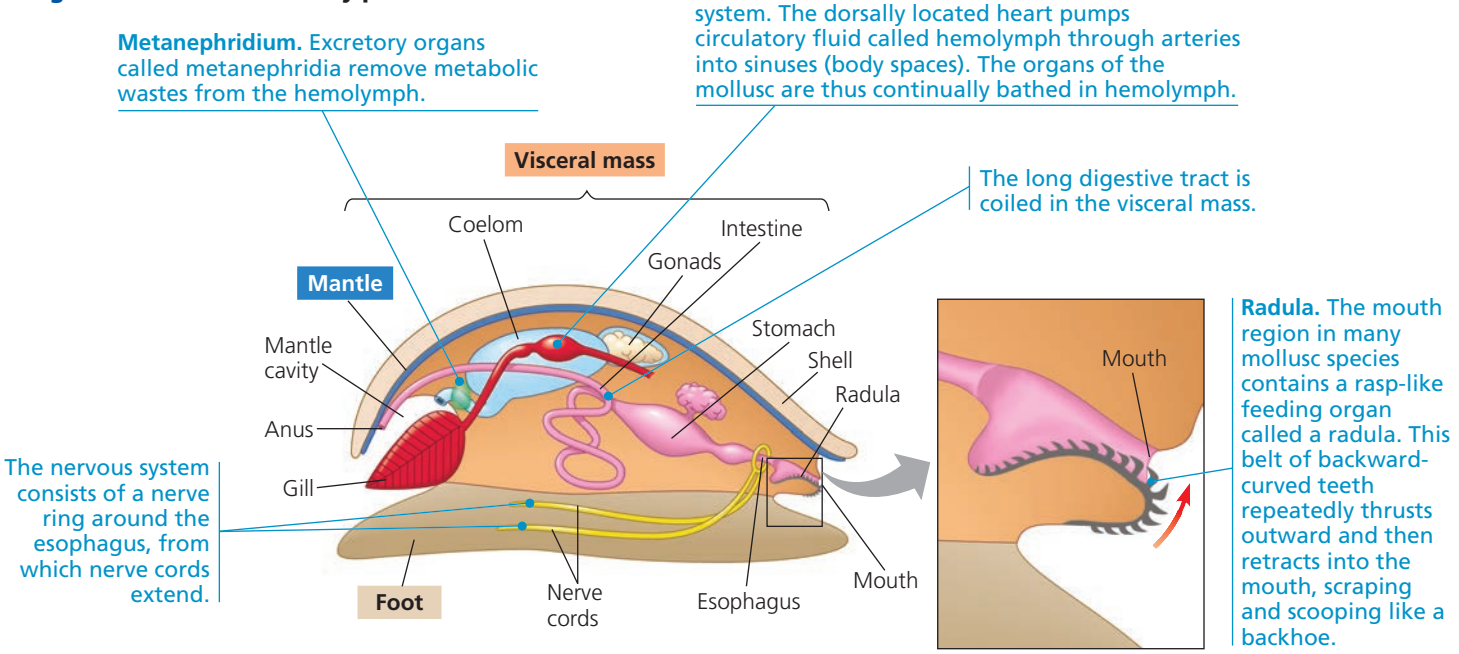
The basic body plan of molluscs has evolved in various ways in the phylum's eight major clades. We'll examine four of those clades here: Polyplacophora (chitons), Gastropoda (snails and slugs), Bivalvia (clams, oysters, and other bivalves), and Cephalopoda (squids, octopuses, cuttlefishes, and chambered nautiluses). We will then focus on threats facing some groups of molluscs.

Chitons

Chitons have an oval-shaped body and a shell composed of eight dorsal plates **(Figure 33.17)**. The chiton's body itself, however, is unsegmented. You can find these marine animals clinging to rocks along the shore during low tide. If you try to dislodge a chiton by hand, you will be surprised at how well its foot, acting as a suction cup, grips the rock. A chiton can also use its foot to creep slowly over the rock surface. Chitons use their radula to scrape algae off the rock surface.

Figure 33.17 A chiton. Note the eight-plate shell characteristic of molluscs in the clade Polyplacophora.

Figure 33.16 The basic body plan of a mollusc.



Gastropods

About three-quarters of all living species of molluscs are gastropods **(Figure 33.18)**. Most gastropods are marine, but there are also freshwater species. Still other gastropods have adapted to life on land, where snails and slugs thrive in habitats ranging from deserts to rain forests.

Gastropods move literally at a snail's pace by a rippling motion of their foot or by means of cilia—a slow process that can leave them vulnerable to attack. Most gastropods have a single, spiraled shell into which the animal can retreat when threatened. The shell, which is secreted by glands at the edge of the mantle, has several functions, including protecting the animal's soft body from injury and dehydration. One of its most important roles is as a defense against predators, as is demonstrated by comparing populations with different histories of predation (see the **Scientific Skills Exercise**). As they move slowly about, most gastropods use their radula to graze on algae or plants. Several groups, however, are predators, and their radula has become modified for boring holes in the shells of other molluscs or for tearing apart prey. In the cone snails, the teeth of the radula act as poison darts that are used to subdue prey.

(a) A land snail

(b) A sea slug. Nudibranchs, or sea slugs, lost their shell during their evolution.

Figure 33.18
Gastropods. The many species of gastropods have colonized terrestrial as well as aquatic environments.

SCIENTIFIC SKILLS EXERCISE

Understanding Experimental Design and Interpreting Data

Is There Evidence of Selection for Defensive Adaptations in Mollusc Populations Exposed to Predators? The fossil record shows that historically, increased risk to prey species from predators is often accompanied by increased incidence and expression of prey defenses. Researchers tested whether populations of the predatory European green crab (*Carcinus maenas*) have exerted similar selective pressures on its gastropod prey, the flat periwinkle (*Littorina obtusata*). Periwinkles from southern sites in the Gulf of Maine have experienced predation by European green crabs for over 100 generations, at about one generation per year. Periwinkles from northern sites in the Gulf have been interacting with the invasive green crabs for relatively few generations, as the invasive crabs spread to the northern Gulf comparatively recently.

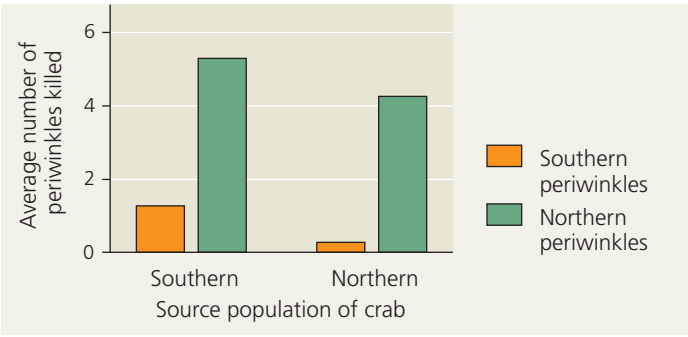
A periwinkle

Previous research shows that (1) flat periwinkle shells recently collected from the Gulf are thicker than those collected in the late 1800s, and (2) periwinkle populations from southern sites in the Gulf have thicker shells than periwinkle populations from northern sites. In this exercise, you'll interpret the design and results of the researchers' experiment studying the rates of predation by crabs on periwinkles from northern and southern populations.

How the Experiment Was Done The researchers collected periwinkles and crabs from sites in the northern and southern Gulf of Maine, separated by 450 km of coastline. A single crab was placed in a cage with eight periwinkles of different sizes. After three days, researchers assessed the fate of the eight periwinkles. Four different treatments were set up, with crabs from northern or southern populations offered periwinkles from northern and southern populations. All crabs were of similar size and included equal numbers of males and females. Each experimental treatment was tested 12 to 14 times.

In a second part of the experiment, the bodies of periwinkles from northern and southern populations were removed from their shells and presented to crabs from northern and southern populations.

Data from the Experiment



Data from R. Rochette et al., Interaction between an invasive decapod and a native gastropod: Predator foraging tactics and prey architectural defenses, *Marine Ecology Progress Series* 330:179–188 (2007).

When the researchers presented the crabs with unshelled periwinkles, all the unshelled periwinkles were consumed in less than an hour.

INTERPRET THE DATA

1. What hypotheses were the researchers testing in this study? What are the independent variables? What are the dependent variables?
2. Why did the research team set up four different treatments?
3. Why did researchers present unshelled periwinkles to the crabs? What do the results of this part of the experiment indicate?
4. Summarize the results of the experiment in words. Do these results support the hypothesis you identified in question 1? Explain.
5. Suggest how natural selection may have affected populations of flat periwinkles in the southern Gulf of Maine over the last 100 years.

Instructors: A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Many gastropods have a head with eyes at the tips of tentacles. Terrestrial snails lack the gills typical of most aquatic gastropods. Instead, the lining of their mantle cavity functions as a lung, exchanging respiratory gases with the air.

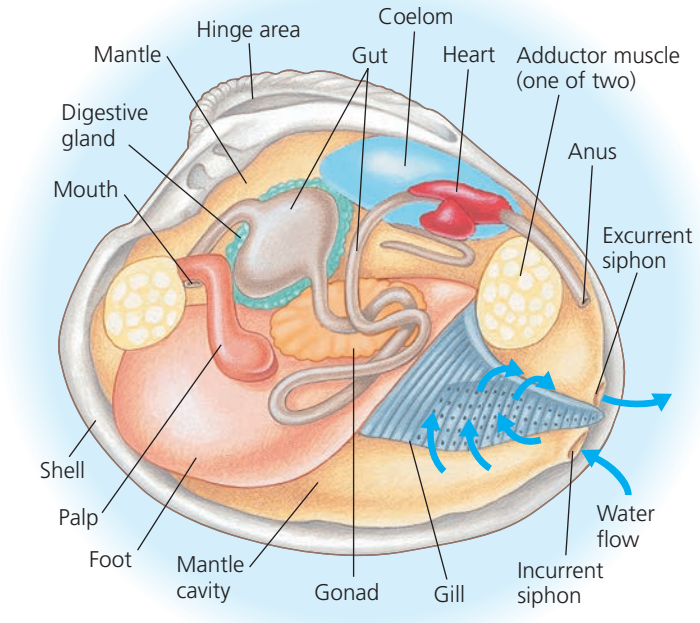
Bivalves

The molluscs of the clade Bivalvia are all aquatic and include many species of clams, oysters, mussels, and scallops. Bivalves have a shell divided into two halves **(Figure 33.19)**. The halves are hinged, and powerful adductor muscles draw them tightly together to protect the animal's soft body. Bivalves have no distinct head, and the radula has been lost. Some bivalves have eyes and sensory tentacles along the outer edge of their mantle.

The mantle cavity of a bivalve contains gills that are used for feeding as well as gas exchange in most species **(Figure 33.20)**.

Figure 33.19 A bivalve. This scallop has many eyes (dark blue spots) lining each half of its hinged shell.

Figure 33.20 Anatomy of a clam. Food particles suspended in water that enters through the incurrent siphon are collected by the gills and passed via cilia and the palps to the mouth.



Most bivalves are suspension feeders. They trap small food particles in mucus that coats their gills, and cilia then convey those particles to the mouth. Water enters the mantle cavity through an incurrent siphon, passes over the gills, and then exits the mantle cavity through an excurrent siphon.

Most bivalves lead sedentary lives, a characteristic suited to suspension feeding. Mussels secrete strong threads that tether them to rocks, docks, boats, and the shells of other animals. However, clams can pull themselves into the sand or mud, using their muscular foot for an anchor, and scallops can skitter along the seafloor by flapping their shells, rather like the mechanical false teeth sold in novelty shops.

Cephalopods

Cephalopods are active marine predators **(Figure 33.21)**. They use their tentacles to grasp prey, which they then bite with beak-like jaws and immobilize with a poison present in their saliva. The foot of a cephalopod has become modified into a muscular excurrent siphon and part of the tentacles. Squids dart about by drawing water into their mantle cavity and then firing a jet of water through the excurrent siphon; they steer by pointing the siphon in different directions. Octopuses use a similar mechanism to escape predators.

The mantle covers the visceral mass of cephalopods, but the shell is generally reduced and internal (in most species) or missing altogether (in some cuttlefishes and some octopuses).

Figure 33.21 Cephalopods.

▶ Squids are speedy carnivores with beak-like jaws and well-developed eyes.

◀ Octopuses are considered among the most intelligent invertebrates.

▶ Chambered nautiluses are the only living cephalopods with an external shell.

One small group of cephalopods with external shells, the chambered nautiluses, survives today.

Cephalopods are the only molluscs with a *closed circulatory system*, in which the blood remains separate from fluid in the body cavity. They also have well-developed sense organs and a complex brain. The ability to learn and behave in a complex manner is probably more critical to fast-moving predators than to sedentary animals such as clams.

The ancestors of octopuses and squids were probably shelled molluscs that took up a predatory lifestyle; the shell was lost in later evolution. Shelled cephalopods called **ammonites**, some of them as large as truck tires, were the dominant invertebrate predators of the seas for hundreds of millions of years until their disappearance during the mass extinction at the end of the Cretaceous period, 65.5 million years ago.

Most species of squid are less than 75 cm long, but some are much larger. The giant squid (*Architeuthis dux*), for example, has an estimated maximum length of 13 m for females and 10 m for males. The colossal squid (*Mesonychoteuthis hamiltoni*), is even larger, with an estimated maximum length of 14 m. Unlike *A. dux*, which has large suckers and small teeth on its tentacles, *M. hamiltoni* has two rows of sharp hooks at the ends of its tentacles that can inflict deadly lacerations.

It is likely that *A. dux* and *M. hamiltoni* spend most of their time in the deep ocean, where they may feed on large fishes. Remains of both giant squid species have been found in the stomachs of sperm whales, which are probably their only natural predator. Scientists first photographed *A. dux* in the wild in 2005 while it was attacking baited hooks at a depth of 900 m. *M. hamiltoni* has yet to be observed in nature. Overall, much remains to be learned about these marine giants.

Protecting Freshwater and Terrestrial Molluscs

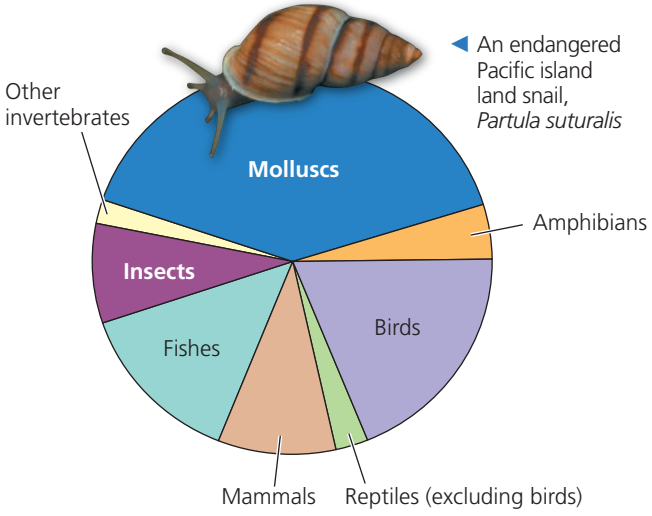
Species extinction rates have increased dramatically in the last 400 years, raising concern that a sixth, human-caused mass extinction may be under way (see Concept 25.4). Among the many taxa under threat, molluscs have the dubious distinction of being the animal group with the largest number of documented extinctions **(Figure 33.22)**.

Threats to molluscs are especially severe in two groups, freshwater bivalves and terrestrial gastropods. For example, the pearl mussels, a group of freshwater bivalves that can make natural pearls (gems that form when a mussel or oyster secretes layers of a lustrous coating around a grain of sand or other small irritant), are among the world's most endangered animals. Roughly 10% of the 300 pearl mussel species that once lived in North America have become extinct in the last 100 years, and over two-thirds of those that remain are threatened by extinction. Terrestrial gastropods, such as the snail in Figure 33.22, are faring no better. Hundreds of Pacific island land snails have disappeared since 1800. Overall, more than 50% of the Pacific island land snails are extinct or under imminent threat of extinction.

Threats faced by freshwater and terrestrial molluscs include habitat loss, pollution, competition or predation by non-native species, and overharvesting by humans. Is it too late to protect these molluscs? In some locations, reducing water pollution and changing how water is released from dams have led to dramatic rebounds in pearl mussel populations. Such results provide hope that with corrective measures, other endangered mollusc species can be revived.

Figure 33.22 The silent extinction. Molluscs account for a largely unheralded but sobering 40% of all documented extinctions of animal species. These extinctions have resulted from habitat loss, pollution, introduced species, overharvesting, and other human actions. Many pearl mussel populations, for example, were driven to extinction by overharvesting for their shells, which were used to make buttons and other goods. Land snails are highly vulnerable to the same threats; like pearl mussels, they are among the world's most imperiled animal groups.

MAKE CONNECTIONS ➤ *Freshwater bivalves feed on and can reduce the abundance of photosynthetic protists and bacteria. As such, would the extinction of freshwater bivalves likely have weak or strong effects on aquatic communities (see Concept 28.6)? Explain.*



▲ Recorded extinctions of animal species

▲ Workers on a mound of pearl mussels killed to make buttons (ca. 1919)

Annelids

Annelida means "little rings," referring to the annelid body's resemblance to a series of fused rings. Annelids are segmented worms that live in the sea, in most freshwater habitats, and in damp soil. Annelids are coelomates, and they range in length from less than 1 mm to more than 3 m.

Traditionally, the phylum Annelida was divided into three main groups, Polychaeta (the polychaetes), Oligochaeta (the oligochaetes), and Hirudinea (the leeches). The names of the first two of these groups reflected the relative number of chaetae, bristles made of chitin, on their bodies: polychaetes (from the Greek *poly*, many, and *chaitē*, long hair) have many more chaetae per segment than do oligochaetes.

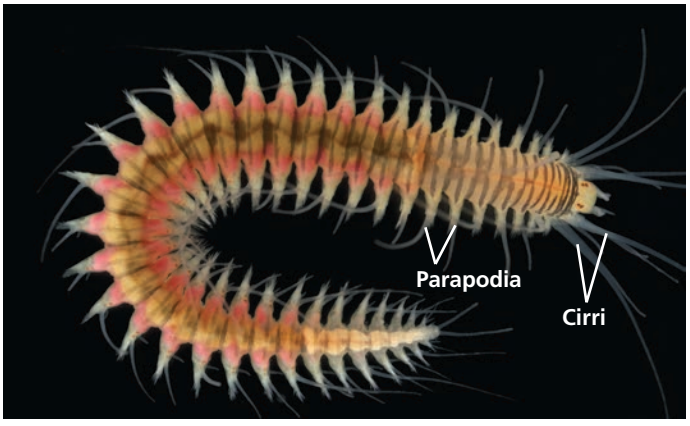
However, a 2011 phylogenomic study and other recent molecular analyses have indicated that the oligochaetes are a subgroup of the polychaetes, making the polychaetes (as defined morphologically) a paraphyletic group. Likewise, the leeches have been shown to be a subgroup of the oligochaetes. As a result, these traditional names are no longer used to describe the evolutionary history of the annelids. Instead, current evidence indicates that the annelids can be divided into two major clades, Errantia and Sedentaria—a grouping that reflects broad differences in lifestyle.

Errantians

Clade Errantia (from the Old French *errant*, traveling) is a large and diverse group, most of whose members are marine **(Figure 33.23)**. As their name suggests, many errantians are mobile; some swim among the plankton (small, drifting organisms), while many others crawl on or burrow in the seafloor. Many are predators, while others are grazers that feed on large, multicellular algae. The group also includes some relatively immobile species, such as the tube-dwelling *Platynereis*, a marine species that recently has become a model organism for studying neurobiology and development.

In many errantians, each body segment has a pair of prominent paddle-like or ridge-like structures called parapodia ("beside feet") that function in locomotion (see Figure 33.23).

Figure 33.23 An errantian, the predator *Nereimyra punctata*. This marine annelid ambushes prey from burrows it has constructed on the seafloor. *N. punctata* hunts by touch, detecting its prey with long sensory organs called cirri that extend from the burrow.



Each parapodium has numerous chaetae. (Possession of parapodia with numerous chaetae is not unique to Errantia, however, as some members of the other major clade of annelids, Sedentaria, also have these features.) In many species, the parapodia are richly supplied with blood vessels and also function as gills. Errantians also tend to have well-developed jaws and sensory organs, as might be expected of predators or grazers that move about in search of food.

Sedentarians

Species in the other major clade of annelids, Sedentaria (from the Latin *sedere*, sit), tend to be less mobile than those in Errantia. Some species burrow slowly through marine sediments or soil, while others live within tubes that protect and support their soft bodies. Tube-dwelling sedentarians often have elaborate gills or tentacles used for filter feeding **(Figure 33.24)**.

Although the Christmas tree worm shown in Figure 33.24 once was classified as a "polychaete," current evidence indicates it is a sedentarian. The clade Sedentaria also contains former "oligochaetes," including the two groups we turn to next, the leeches and the earthworms.

Leeches Some leeches are parasites that suck blood by attaching temporarily to other animals, including humans **(Figure 33.25)**, but most are predators that feed on other

Figure 33.24 The Christmas tree worm, *Spirobranchus giganteus*. This sedentarian's two tree-shaped whorls are tentacles, which it uses in gas exchange and to collect food particles from the water. The tentacles emerge from a calcium carbonate tube secreted by the worm that protects and supports its soft body.

Figure 33.25 A leech. A nurse applied this medicinal leech (*Hirudo medicinalis*) to a patient's sore thumb to drain blood from a hematoma (an abnormal accumulation of blood around an internal injury).

invertebrates. Leeches range in length from 1 to 30 cm. Most leeches inhabit fresh water, but there are also marine species and terrestrial leeches, which live in moist vegetation. Some parasitic species use bladelike jaws to slit the skin of their host. The host is usually oblivious to this attack because the leech secretes an anesthetic. After making the incision, the leech secretes a chemical, hirudin, which keeps the blood of the host from coagulating near the incision. The parasite then sucks as much blood as it can hold, often more than ten times its own weight. After this gorging, a leech can last for months without another meal.

Until the 20th century, leeches were frequently used for bloodletting. Today they are used to drain blood that accumulates in tissues following certain injuries or surgeries. In addition, forms of hirudin produced with recombinant DNA techniques can be used to dissolve unwanted blood clots that form during surgery or as a result of heart disease.

Earthworms Earthworms eat their way through the soil, extracting nutrients as the soil passes through the alimentary canal. Undigested material, mixed with mucus secreted into the canal, is eliminated as fecal castings through the anus. Farmers value earthworms because the animals till and aerate the earth, and their castings improve the texture of the soil. (Charles Darwin estimated that one acre of farmland contains about 50,000 earthworms, producing 18 tons of castings per year.)

A guided tour of the anatomy of an earthworm, which is representative of annelids, is shown in **Figure 33.26**.

Giant Australian earthworm

Figure 33.26 Anatomy of an earthworm, a sedentarian.

Earthworms are hermaphrodites, but they do cross-fertilize. Two earthworms mate by aligning themselves in opposite directions in such a way that they exchange sperm, and then they separate. Some earthworms can also reproduce asexually by fragmentation followed by regeneration.

As a group, Lophotrochozoa encompasses a remarkable range of body plans, as illustrated by members of such phyla as Syndermata, Ectoprocta, Mollusca, and Annelida. Next we'll explore the diversity of Ecdysozoa, a dominant presence on Earth in terms of sheer number of species.

CONCEPT CHECK 33.3

1. Explain how tapeworms can survive without a coelom, a mouth, a digestive system, or an excretory system.
2. Annelid anatomy can be described as "a tube within a tube." Explain.
3. **MAKE CONNECTIONS** ➤ Explain how the molluscan foot in gastropods and the excurrent siphon in cephalopods represent examples of descent with modification (see Concept 21.2).

For suggested answers, see Appendix A.

CONCEPT 33.4

Ecdysozoans are the most species-rich animal group

Although defined primarily by molecular evidence, the clade Ecdysozoa includes animals that shed a tough external coat (**cuticle**) as they grow; in fact, the group derives its name from this process, which is called *ecdysis*, or **molting**. Ecdysozoa includes about eight animal phyla and contains more known species than all other animal, protist, fungus, and plant groups combined. Here we'll focus on the two largest ecdysozoan phyla, the nematodes and arthropods, which are among the most successful and abundant of all animal groups.

Nematodes

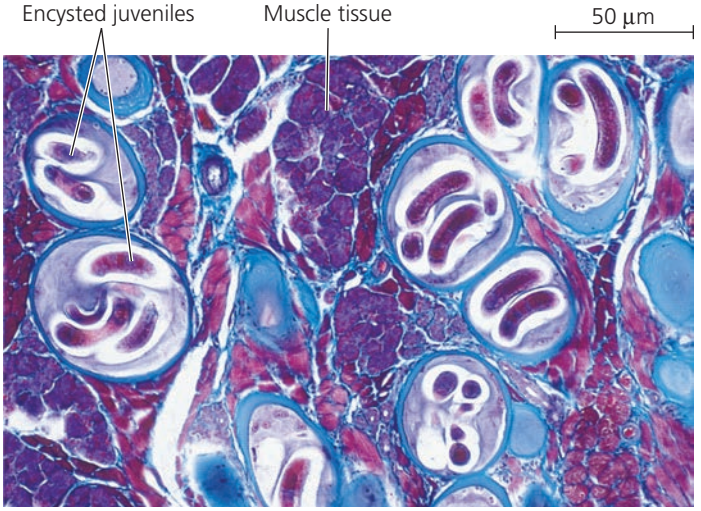
Among the most ubiquitous of animals, nematodes (phylum Nematoda), or roundworms, are found in most aquatic habitats, in the soil, in the moist tissues of plants, and in the body fluids and tissues of animals. The cylindrical bodies of nematodes range from less than 1 mm to more than 1 m long, often tapering to a fine tip at the posterior end and to a blunter tip at the anterior end **(Figure 33.27)**. A nematode's body is covered by a tough cuticle (a type of exoskeleton); as the worm grows, it periodically sheds its old cuticle and secretes a new, larger one. Nematodes have an alimentary canal, though they lack a circulatory system. Nutrients are transported throughout the body via fluid in the pseudocoelom. The body wall muscles are all longitudinal, and their contraction produces a thrashing motion.

➤ Figure 33.27 A free-living nematode. (colorized SEM).

Multitudes of nematodes live in moist soil and in decomposing organic matter on the bottoms of lakes and oceans. While 25,000 species are known, perhaps 20 times that number actually exist. It has been said that if nothing of Earth or its organisms remained but nematodes, they would still preserve the outline of the planet and many of its features. These free-living worms play an important role in decomposition and nutrient cycling, but little is known about most species. One species of soil nematode, *Caenorhabditis elegans*, however, is very well studied and has become a model research organism in biology (see Concept 46.3). Ongoing studies of *C. elegans* are providing insight into mechanisms involved in aging in humans, as well as many other topics.

Phylum Nematoda includes many species that parasitize plants, and some are major agricultural pests that attack the roots of crops. Other nematodes parasitize animals. Some of these species benefit humans by attacking insects such as cutworms that feed on the roots of crop plants. On the other hand, humans are hosts to at least 50 nematode species, including various pinworms and hookworms. One notorious nematode is *Trichinella spiralis*, the worm that causes trichinosis **(Figure 33.28)**. Humans acquire this nematode

➤ Figure 33.28 Juveniles of the parasitic nematode *Trichinella spiralis* encysted in human muscle tissue. (LM).



by eating raw or undercooked pork or other meat (including wild game such as bear or walrus) that has juvenile worms encysted in the muscle tissue. Within the human intestines, the juveniles develop into sexually mature adults. Females burrow into the intestinal muscles and produce more juveniles, which bore through the body or travel in lymphatic vessels to other organs, including skeletal muscles, where they encyst.

Parasitic nematodes have an extraordinary molecular toolkit that enables them to redirect some of the cellular functions of their hosts. Some species inject their plant hosts with molecules that induce the development of root cells, which then supply nutrients to the parasites. When *Trichinella* parasitizes animals, it regulates the expression of specific muscle cell genes encoding proteins that make the cell elastic enough to house the nematode. Additionally, the infected muscle cell releases signals that promote the growth of new blood vessels, which then supply the nematode with nutrients.

Arthropods

Zoologists estimate that there are about a billion billion (10^{18}) arthropods living on Earth. More than 1 million arthropod species have been described, most of which are insects. In fact, two out of every three known species are arthropods, and members of the phylum Arthropoda can be found in nearly all habitats of the biosphere. By the criteria of species diversity, distribution, and sheer numbers, arthropods must be regarded as the most successful of all animal phyla.

Arthropod Origins

Biologists hypothesize that the diversity and success of **arthropods** are related to their body plan—their segmented body, hard exoskeleton, and jointed appendages. How did this body plan arise and what advantages did it provide?

The earliest fossils of arthropods are from the Cambrian explosion (535–525 million years ago), indicating that the arthropods are at least that old. The fossil record of the Cambrian explosion also contains many species of *lobopods*, a group from which arthropods may have evolved. Lobopods such as *Hallucigenia* (see Figure 32.7) had segmented bodies, but most of their body segments were identical to one another. Early arthropods, such as the trilobites, also showed little variation from segment to segment **(Figure 33.29)**. As arthropods continued to evolve, groups of segments tended to become functionally united into "body regions" specialized for tasks such as feeding, walking, or swimming. These evolutionary changes resulted not only in great diversification but also in efficient body plans that permit the division of labor among different body regions.

➤ Figure 33.29 A trilobite fossil. Trilobites were common denizens of the shallow seas throughout the Paleozoic era but disappeared with the great Permian extinctions about 250 million years ago. Paleontologists have described about 4,000 trilobite species.

What genetic changes led to the increasing complexity of the arthropod body plan? Arthropods today have two unusual *Hox* genes, both of which influence segmentation. To test whether these genes could have driven the evolution of increased body segment diversity in arthropods, researchers studied *Hox* genes in onychophorans (see Figure 33.3), close relatives of arthropods **(Figure 33.30)**. Their results indicate that the diversity of arthropod body plans did *not* arise from the acquisition of new *Hox* genes. Instead, the evolution of body segment diversity in arthropods was probably driven by changes in the sequence or regulation of existing *Hox* genes (see Concept 25.5).

General Characteristics of Arthropods

Over the course of evolution, the appendages of some arthropods have become modified, specializing in functions such as walking, feeding, sensory reception, reproduction, and defense. Like the appendages from which they were derived, these modified structures are jointed and come in pairs. **Figure 33.31** illustrates the diverse appendages and other arthropod characteristics of a lobster.

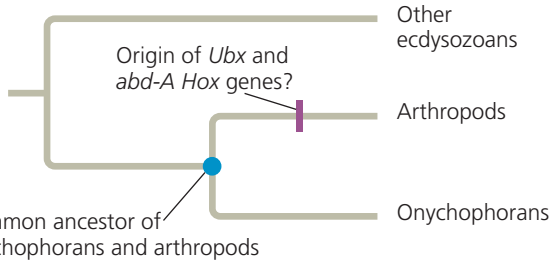
The body of an arthropod is completely covered by the cuticle, an exoskeleton constructed from layers of protein and the polysaccharide chitin. As you know if you've ever eaten a crab or lobster, the cuticle can be thick and hard over some parts of the body and thin and flexible over others, such as the joints. The rigid exoskeleton protects the animal and provides points of attachment for the muscles that move the appendages. But it also prevents the arthropod from growing, unless it occasionally sheds its exoskeleton and produces a larger one. This molting process is energetically expensive, and it leaves the arthropod vulnerable to predation and other dangers until its new, soft exoskeleton hardens.

When the arthropod exoskeleton first evolved in the sea, its main functions were probably protection and anchorage for muscles, but it later enabled certain arthropods to live on land. The exoskeleton's relative impermeability to water helped prevent desiccation, and its strength provided support when arthropods left the buoyancy of water. Fossil

Figure 33.30

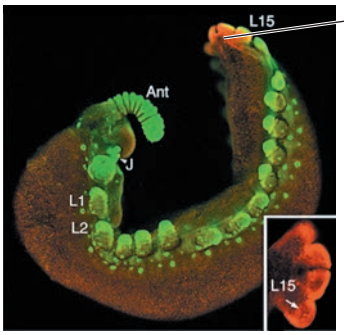
Inquiry Did the arthropod body plan result from new *Hox* genes?

Experiment One hypothesis suggests that the arthropod body plan resulted from the origin (by gene duplication and subsequent mutations) of two unusual *Hox* genes found in arthropods: *Ultrabithorax* (*Ubx*) and *abdominal-A* (*abd-A*). Researchers tested this hypothesis using onychophorans, a group of invertebrates closely related to arthropods. Unlike many living arthropods, onychophorans have a body plan in which most body segments are identical to one another. If the origin of the *Ubx* and *abd-A Hox* genes drove the evolution of body segment diversity in arthropods, these genes probably arose on the arthropod branch of the evolutionary tree:



According to this hypothesis, *Ubx* and *abd-A* would not have been present in the common ancestor of arthropods and onychophorans; hence, onychophorans should not have these genes. The researchers examined the *Hox* genes of the onychophoran *Acanthokara kaputensis*.

Results The onychophoran *A. kaputensis* has all arthropod *Hox* genes, including *Ubx* and *abd-A*.



Red indicates the body regions of this onychophoran embryo in which *Ubx* or *abd-A* genes were expressed. (The inset shows this area enlarged.)

Ant = antenna
J = jaws
L1–L15 = body segments

Conclusion The evolution of increased body segment diversity in arthropods was not related to the origin of new *Hox* genes.

Data from J. K. Grenier et al., Evolution of the entire arthropod *Hox* gene set predated the origin and radiation of the onychophoran/arthropod clade, *Current Biology* 7:547–553 (1997).

WHAT IF? ➤ *Suppose* A. kaputensis *did* not *have the* Ubx *and* abd-A Hox *genes. How would the conclusions of this study have been affected? Explain.*

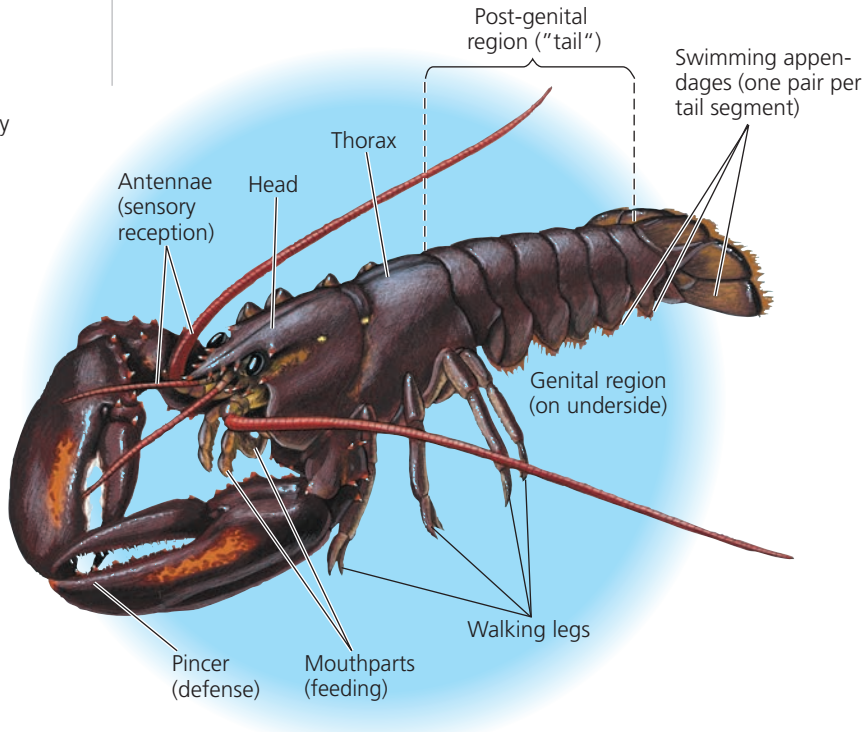
Figure 33.31 External anatomy of an arthropod. Many of the distinctive features of arthropods are apparent in this dorsal view of a lobster. The body is segmented, but this character is obvious only in the post-genital region or "tail," located behind the genitals. The appendages (including antennae, pincers, mouthparts, walking legs, and swimming appendages) are jointed. The head bears a pair of compound (multilens) eyes. The whole body, including appendages, is covered by an exoskeleton.

evidence suggests that arthropods were among the first animals to colonize land, roughly 450 million years ago. These fossils include fragments of arthropod remains, as well as possible millipede burrows. Arthropod fossils from several continents indicate that by 410 million years ago, millipedes, centipedes, spiders, and a variety of wingless insects all had colonized land.

Arthropods have well-developed sensory organs, including eyes, olfactory (smell) receptors, and antennae that function in both touch and smell. Most sensory organs are concentrated at the anterior end of the animal, although there are interesting exceptions. Female butterflies, for example, "taste" plants using sensory organs on their feet.

Like many molluscs, arthropods have an **open circulatory system**, in which fluid called *hemolymph* is propelled by a heart through short arteries and then into spaces called sinuses surrounding the tissues and organs. (The term *blood* is generally reserved for fluid in a closed circulatory system.) Hemolymph reenters the arthropod heart through pores that are usually equipped with valves. The hemolymph-filled body sinuses are collectively called the *hemocoel*, which is not part of the coelom. Although arthropods are coelomates, in most species the coelom that forms in the embryo becomes much reduced as development progresses, and the hemocoel becomes the main body cavity in adults.

A variety of specialized gas exchange organs have evolved in arthropods. These organs allow the diffusion of respiratory gases in spite of the exoskeleton. Most aquatic species have gills with thin, feathery extensions that expose a large surface area to the surrounding water. Terrestrial arthropods generally have internal surfaces specialized for gas exchange.



For example, most insects have tracheal systems, branched air ducts leading into the interior of the body from pores in the cuticle.

Morphological and molecular data suggest that living arthropods consist of three major lineages that diverged early in the phylum's evolution: **chelicerates** (sea spiders, horseshoe crabs, scorpions, ticks, mites, and spiders); **myriapods** (centipedes and millipedes); and **pancrustaceans** (a recently defined, diverse group that includes insects as well as lobsters, shrimp, barnacles, and other crustaceans).

Chelicerates

Chelicerates (clade Chelicerata) are named for clawlike feeding appendages called **chelicerae**, which serve as pincers or fangs. Chelicerates lack antennae, and most have simple eyes (eyes with a single lens).

The earliest chelicerates were **eurypterids**, or water scorpions. These marine and freshwater predators grew up to 3 m long; it is thought that some species could have walked on land, much as land crabs do today. Most of the marine chelicerates, including all of the eurypterids, are extinct. Among the marine chelicerates that survive today are the sea spiders (pycnogonids) and horseshoe crabs **(Figure 33.32)**.

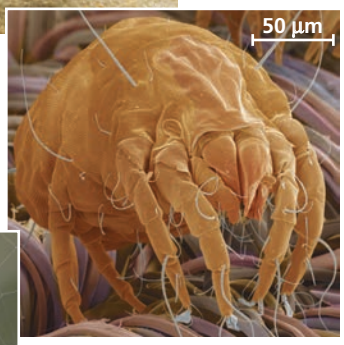
The bulk of modern chelicerates are **arachnids**, a group that includes scorpions, spiders, ticks, and mites **(Figure 33.33)**. Nearly all ticks are bloodsucking parasites that live on the body surfaces of reptiles or mammals. Parasitic mites live on or in a wide variety of vertebrates, invertebrates, and plants.

Arachnids have six pairs of appendages: the chelicerae; a pair of appendages called *pedipalps* that function in sensing, feeding, defense, or reproduction; and four pairs of walking legs. Spiders use their fang-like chelicerae, which are equipped with poison glands, to attack prey. As the chelicerae pierce the prey, the spider secretes digestive juices onto the prey's torn tissues. The food softens, and the spider sucks up the liquid meal. In most spiders, gas exchange is carried out by **book lungs**, stacked platelike structures contained in an internal chamber **(Figure 33.34)**. The extensive surface area of these respiratory organs enhances the exchange of O_2 and CO_2 between the hemolymph and air.

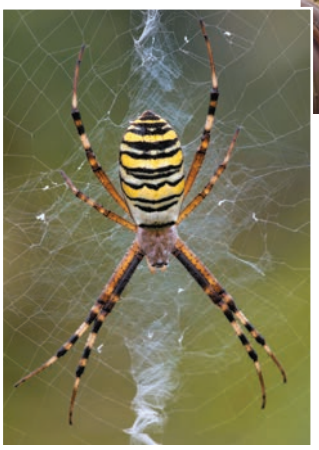
A unique adaptation of many spiders is the ability to catch insects by constructing webs of silk, a liquid protein produced by specialized abdominal glands. The silk is spun by organs called spinnerets into fibers that then solidify. Each spider engineers a web characteristic of its species and builds it perfectly on the first try, indicating that this complex behavior is inherited. Various spiders also use silk in other ways: as droplines for rapid escape, as a cover for eggs, and even as "gift wrap" for food that males offer females during courtship. Many

Figure 33.32 Horseshoe crabs (*Limulus polyphemus*). Common on the Atlantic and Gulf coasts of the United States, these "living fossils" have changed little in hundreds of millions of years. They are surviving members of a rich diversity of chelicerates that once filled the seas.

▲ Scorpions have pedipalps that are pincers specialized for defense and the capture of food. The tip of the tail bears a poisonous stinger.



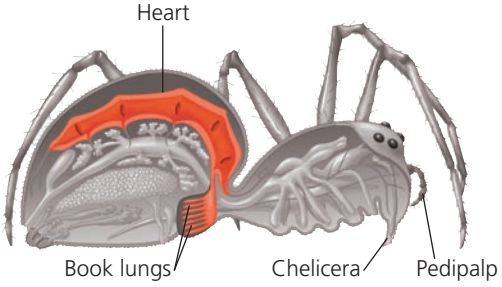
▲ Dust mites are ubiquitous scavengers in human dwellings but are harmless except to those people who are allergic to them (colorized SEM).



◀ Web-building spiders are generally most active during the daytime.

Figure 33.33 Arachnids.

Figure 33.34 Book lungs.



(a) Millipede

(b) Centipede

Figure 33.35 Myriapods.

small spiders also extrude silk into the air and let themselves be transported by wind, a behavior known as "ballooning."

Myriapods

Millipedes and centipedes belong to the clade Myriapoda **(Figure 33.35)**. All living myriapods are terrestrial. The myriapod head has a pair of antennae and three pairs of appendages modified as mouthparts, including the jaw-like mandibles.

Millipedes have a large number of legs, though fewer than the thousand their name implies. Each trunk segment is formed from two fused segments and bears two pairs of legs (see Figure 33.35a). Millipedes eat decaying leaves and other plant matter. They may have been among the earliest animals on land, living on mosses and early vascular plants.

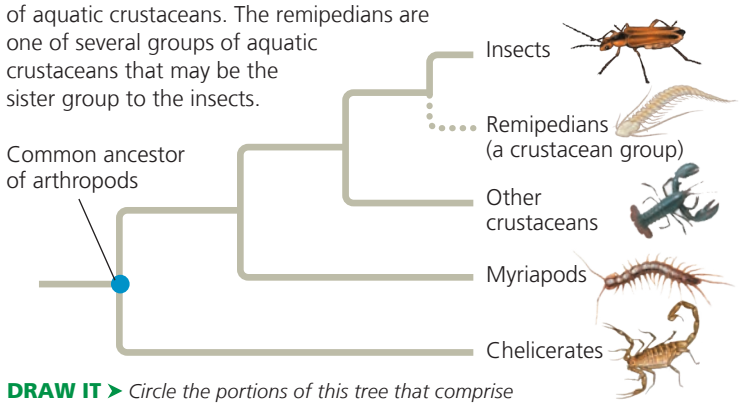
Unlike millipedes, centipedes are carnivores. Each segment of a centipede's trunk region has one pair of legs (see Figure 33.35b). Centipedes have poison claws on their foremost trunk segment that paralyze prey and aid in defense.

Pancrustaceans

A series of recent papers, including a 2010 phylogenomic study, present evidence that terrestrial insects are more closely related to lobsters and other crustaceans than they are to the terrestrial group we just discussed, the myriapods (millipedes and centipedes). These studies also suggest that the diverse group of organisms referred to as crustaceans are paraphyletic: Some lineages of crustaceans are more closely related to insects than they are to other crustaceans **(Figure 33.36)**. However, together the insects and crustaceans form a clade, which systematists have named Pancrustacea (from the Greek *pan*, all). We turn next to a description of the members of Pancrustacea, focusing first on crustaceans and then on the insects.

Crustaceans Crustaceans (crabs, lobsters, shrimps, barnacles, and many others) thrive in a broad range of marine, freshwater, and terrestrial environments. Many crustaceans have highly specialized appendages. Lobsters and crayfishes, for instance, have a toolkit of 19 pairs of appendages (see Figure 33.31). The anterior-most appendages form two pairs of antennae; crustaceans are the only arthropods with two pairs. Three or more pairs of appendages are modified as

Figure 33.36 The phylogenetic position of the insects.
Recent results have shown that the insects are nested within lineages of aquatic crustaceans. The remipedians are one of several groups of aquatic crustaceans that may be the sister group to the insects.



DRAW IT *Circle the portions of this tree that comprise the clade Pancrustacea.*

mouthparts, including the hard mandibles. Walking legs are present on the thorax, and, unlike their terrestrial relatives, the insects, crustaceans also have appendages on their postgenital region, or "tail."

Small crustaceans exchange gases across thin areas of the cuticle; larger species have gills. Nitrogenous wastes also diffuse through thin areas of the cuticle, but a pair of glands regulates the salt balance of the hemolymph.

Sexes are separate in most crustaceans. In the case of lobsters and crayfishes, the male uses a specialized pair of abdominal appendages to transfer sperm to the reproductive pore of the female during copulation. Most aquatic crustaceans go through one or more swimming larval stages.

One of the largest groups of crustaceans (numbering over 11,000 species) is the *isopods*, which include terrestrial, freshwater, and marine species. Some isopod species are abundant in habitats at the bottom of the deep ocean. Among the terrestrial isopods are the pill bugs, or wood lice, common on the undersides of moist logs and leaves.

Lobsters, crayfishes, crabs, and shrimps are all relatively large crustaceans called *decapods* **(Figure 33.37)**. The cuticle of decapods is hardened by calcium carbonate. Most decapod species are marine. Crayfishes, however, live in fresh water, and some tropical crabs live on land.

Figure 33.37 A ghost crab, an example of a decapod.
Ghost crabs live on sandy ocean beaches worldwide. Primarily nocturnal, they take shelter in burrows during the day.

Figure 33.38 Krill. These planktonic crustaceans are consumed in vast quantities by some whales.

Figure 33.39 Barnacles. The jointed appendages projecting from the barnacles' shells capture organisms and organic particles suspended in the water.

Many small crustaceans are important members of marine and freshwater plankton communities. Planktonic crustaceans include many species of *copepods*, which are among the most numerous of all animals. Some copepods are grazers that feed upon algae, while others are predators that eat small animals (including smaller copepods!). Copepods are rivaled in abundance by the shrimplike krill, which grow to about 5 cm long **(Figure 33.38)**. A major food source for baleen whales (including blue whales, humpbacks, and right whales), krill are now being harvested in great numbers by humans for food and agricultural fertilizer. The larvae of many larger-bodied crustaceans are also planktonic.

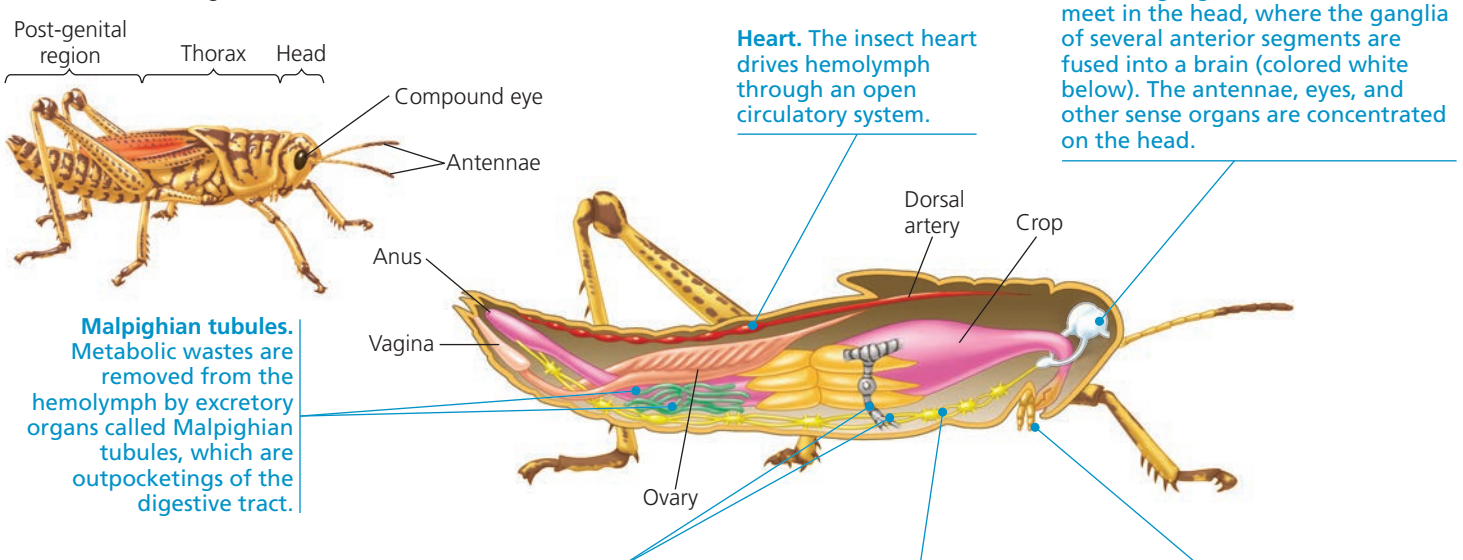
With the exception of a few parasitic species, barnacles are a group of sessile crustaceans whose cuticle is hardened into a shell containing calcium carbonate **(Figure 33.39)**. Most barnacles anchor themselves to rocks, boat hulls, pilings, and other submerged surfaces. Their natural adhesive is as strong as synthetic glues. These barnacles feed by extending appendages from their shell to strain food from the water. Barnacles were not recognized as crustaceans until the 1800s, when naturalists discovered that barnacle larvae resemble the larvae of other crustaceans. The remarkable mix of unique traits and crustacean homologies found in barnacles was a major inspiration to Charles Darwin as he developed his theory of evolution.

We turn now to a group nested within the paraphyletic crustaceans, the insects.

Insects Insects and their six-legged terrestrial relatives form an enormous clade, Hexapoda; we'll focus here on the insects, since as a group they have more described species than all other eukaryotic groups combined. Insects live in almost every terrestrial habitat and in fresh water, and flying insects fill the air. Insects are rare, though not absent, in marine habitats. The internal anatomy of an insect includes several complex organ systems, which are highlighted in **Figure 33.40**.

The oldest insect fossils date to about 415 million years ago. Later, an explosion in insect diversity took place when insect flight evolved during the Carboniferous and Permian periods (359–252 million years ago). An animal that can fly can escape predators, find food and mates, and disperse to new habitats more effectively than an animal that must crawl about on the

Figure 33.40 Anatomy of a grasshopper, an insect. The insect body has three regions: head, thorax, and post-genital region. The segmentation of the thorax and post-genital region is obvious, but the segments that form the head are fused.



ground. Many insects have one or two pairs of wings that emerge from the dorsal side of the thorax. Because the wings are extensions of the cuticle, insects can fly without sacrificing any walking legs **(Figure 33.41)**. By contrast, the flying vertebrates—birds and bats—have one of their two pairs of walking legs modified into wings, making some of these species clumsy on the ground.

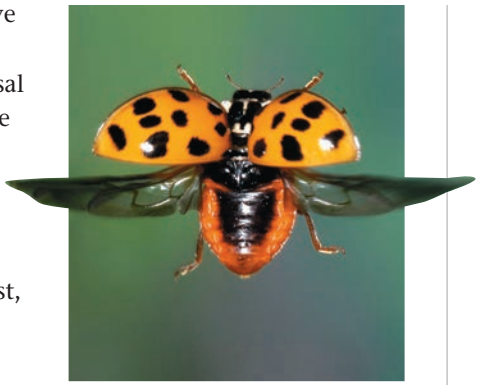


Figure 33.41 Ladybird beetle in flight.

Insects also radiated in response to the origin of new plant species, which provided new sources of food. By the speciation mechanisms described in Concept 24.2, an insect population feeding on a new plant species can diverge from other populations, eventually forming a new species of insect. A fossil record of diverse insect mouthparts, for example, suggests that specialized modes of feeding on gymnosperms and other Carboniferous plants contributed to early adaptive radiations of insects. Later, a major increase in insect diversity appears to have been stimulated by the evolutionary expansion of flowering plants during the mid-Cretaceous period (about 100 million years ago). Although insect and plant diversity decreased during the Cretaceous mass extinction, both groups have rebounded over the past 66 million years. Increases in the diversity of particular insect groups have often been associated with radiations of the flowering plants on which they fed.

Many insects undergo metamorphosis during their development. In the **incomplete metamorphosis** of grasshoppers and some other insect groups, the young (called nymphs) resemble adults but are smaller, have different body proportions, and lack wings. The nymph undergoes a series of molts, each time looking more like an adult. With the final molt, the insect reaches full size, acquires wings, and becomes sexually mature. Insects with **complete metamorphosis** have larval stages specialized for eating and growing that are known by such names as caterpillar, maggot, or grub. The larval stage looks entirely different from the adult stage, which is specialized for dispersal and reproduction. Metamorphosis from the larval stage to the adult occurs during a pupal stage **(Figure 33.42)**.

Reproduction in insects is usually sexual, with separate male and female individuals. Adults come together and recognize each other as members of the same species by advertising with bright colors (as in butterflies), sounds (as in crickets), or odors (as in moths). Fertilization is generally internal. In most species, sperm are deposited directly into the female's vagina at the time of copulation, though in some species the male deposits a sperm packet outside the female, and the female picks it up. An internal structure in the female called the spermatheca stores the sperm, usually enough to fertilize more than one batch of eggs. Many insects mate only once in a lifetime. After mating, a female often lays her eggs on an appropriate food source where the next generation can begin eating as soon as it hatches.

Insects are classified in more than 30 orders, 8 of which are introduced in **Figure 33.43**.

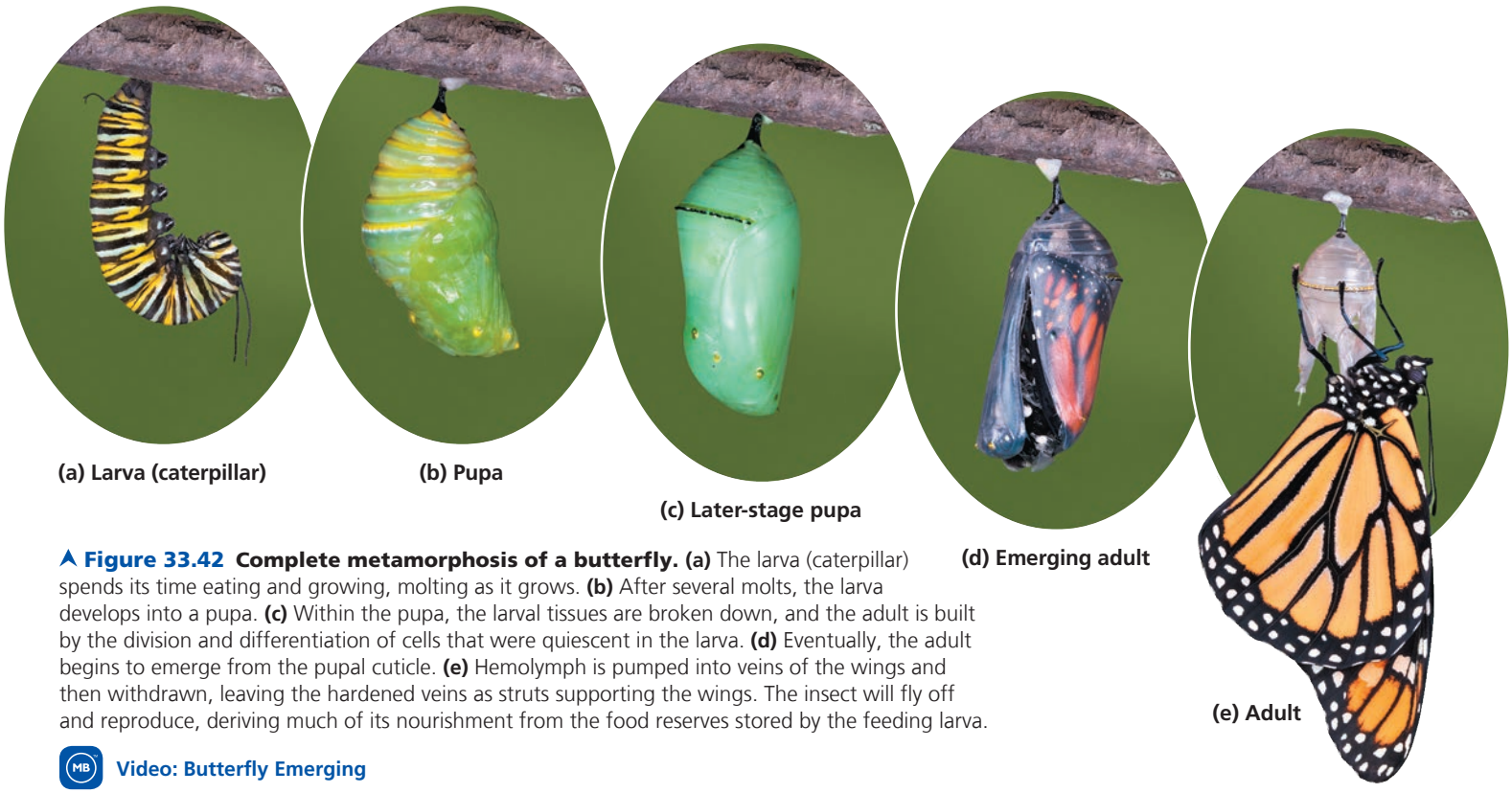


Figure 33.42 Complete metamorphosis of a butterfly. **(a)** The larva (caterpillar) spends its time eating and growing, molting as it grows. **(b)** After several molts, the larva develops into a pupa. **(c)** Within the pupa, the larval tissues are broken down, and the adult is built by the division and differentiation of cells that were quiescent in the larva. **(d)** Eventually, the adult begins to emerge from the pupal cuticle. **(e)** Hemolymph is pumped into veins of the wings and then withdrawn, leaving the hardened veins as struts supporting the wings. The insect will fly off and reproduce, deriving much of its nourishment from the food reserves stored by the feeding larva.

MB **Video: Butterfly Emerging**

Figure 33.43 Exploring Insect Diversity

Although there are more than 30 orders of insects, we'll focus on just 8 here. Two orders of wingless insects, the bristletails (Archaeognatha) and silverfish (Zygentoma), diverged from other insects early in insect evolution. Evolutionary relationships among the other groups discussed here are under debate and so are not depicted on the tree.

Archaeognatha (bristletails; 350 species)

These wingless insects are found under bark and in other moist, dark habitats such as leaf litter, compost piles, and rock crevices. They feed on algae, plant debris, and lichens.

Zygentoma (silverfish; 450 species)

These small, wingless insects have a flattened body and reduced eyes. They live in leaf litter or under bark. They can also infest buildings, where they can become pests.

Winged insects (many orders; six are shown below)

Complete metamorphosis

Coleoptera (beetles; 350,000 species)

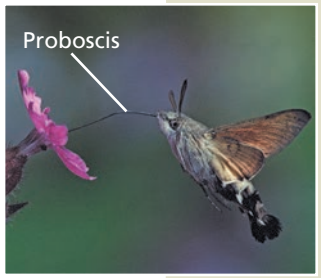
Beetles, such as this male snout weevil (*Rhiastus lasternus*), constitute the most species-rich order of insects. They have two pairs of wings, one of which is thick and stiff, the other membranous. They have an armored exoskeleton and mouthparts adapted for biting and chewing.

Diptera (151,000 species)

Dipterans have one pair of wings; the second pair has become modified into balancing organs called halteres. Their mouthparts are adapted for sucking, piercing, or lapping. Flies and mosquitoes are among the best-known dipterans, which live as scavengers, predators, and parasites. Like many other insects, flies such as this red tachinid (*Adejeania vexatrix*) have well-developed compound eyes that provide a wide-angle view and excel at detecting fast movements.

Hymenoptera (125,000 species)

Most hymenopterans, which include ants, bees, and wasps, are highly social insects. They have two pairs of membranous wings, a mobile head, and chewing or sucking mouthparts. The females of many species have a posterior stinging organ. Many species, such as this European paper wasp (*Polistes dominulus*), build elaborate nests.



Lepidoptera (120,000 species)

Butterflies and moths have two pairs of wings covered with tiny scales. To feed, they uncoil a long proboscis, visible in this photograph of a hummingbird hawkmoth (*Macroglossum stellatarum*). This moth's name refers to its ability to hover in the air while feeding from a flower. Most lepidopterans feed on nectar, but some species feed on other substances, including animal blood or tears.

Incomplete metamorphosis

Hemiptera (85,000 species)

Hemipterans include so-called "true bugs," such as stink bugs, bed bugs, and assassin bugs. (Insects in other orders are sometimes erroneously called bugs.) Hemipterans have two pairs of wings, one pair partly leathery, the other pair membranous. They have piercing or sucking mouthparts and undergo incomplete metamorphosis, as shown in this image of an adult stink bug guarding its offspring (nymphs).

Orthoptera (13,000 species)

Grasshoppers, crickets, and their relatives are mostly herbivorous. They have large hind legs adapted for jumping, two pairs of wings (one leathery, one membranous), and biting or chewing mouthparts. This aptly named spiny devil katydid (*Panacanthus cuspidatus*) has a face and legs specialized for making a threatening display. Male orthopterans commonly make courtship sounds by rubbing together body parts, such as ridges on their hind legs.

Animals as numerous, diverse, and widespread as insects are bound to affect the lives of most other terrestrial organisms, including humans. Insects consume enormous quantities of plant matter; play key roles as predators, parasites, and decomposers; and are an essential source of food for larger animals such as lizards, rodents, and birds. Humans depend on bees, flies, and many other insects to pollinate crops and orchards. In addition, people in many parts of the world eat insects as an important source of protein. On the other hand, insects are carriers for many diseases, including African sleeping sickness (spread by tsetse flies that carry the protist *Trypanosoma*; see Figure 28.7) and malaria (spread by mosquitoes that carry the protist *Plasmodium*; see Figure 23.18 and Figure 28.16).

Insects also compete with humans for food. In parts of Africa, for instance, insects claim about 75% of the crops. In the United States, billions of dollars are spent each year on pesticides, spraying crops with massive doses of some of the deadliest poisons ever invented. Try as they may, not even humans have challenged the preeminence of insects and their arthropod kin. As one prominent entomologist put it: "Bugs are not going to inherit the Earth. They own it now. So we might as well make peace with the landlord."

CONCEPT CHECK 33.4

1. How do nematode and annelid body plans differ?
2. How have modifications in the feeding appendages of arthropods helped them thrive successfully on Earth?
3. **MAKE CONNECTIONS ➤** Historically, annelids and arthropods were viewed as closely related because both have body segmentation. Yet DNA sequence data indicate that annelids belong to one clade (Lophotrochozoa) and arthropods to another (Ecdysozoa). Could traditional and molecular hypotheses be tested by studying the *Hox* genes that control body segmentation (see Concept 20.6)? Explain.

For suggested answers, see Appendix A.

CONCEPT 33.5

Echinoderms and chordates are deuterostomes

Sea stars, sea urchins, and other echinoderms (phylum Echinodermata) may seem to have little in common with vertebrates (animals that have a backbone) and other members of phylum Chordata. Nevertheless, DNA evidence indicates that echinoderms and chordates are closely related, with both phyla belonging to the Deuterostomia clade of bilaterian animals. Echinoderms and chordates also share features characteristic of a deuterostome mode of development, such as radial cleavage and formation of the anus from the blastopore (see Figure 32.10). As discussed in Concept 32.4, however, some animal phyla with members that have deuterostome developmental features, including ectoprocts and brachiopods, are not in the deuterostome clade. Hence, despite its name, the clade Deuterostomia is defined primarily by DNA similarities, not developmental similarities.

Echinoderms

Sea stars (commonly called starfish) and most other groups of **echinoderms** (from the Greek *echin*, spiny, and *derma*, skin) are slow-moving or sessile marine animals. Echinoderms are coelomates. A thin epidermis covers an endoskeleton of hard calcareous plates, and most species are prickly from skeletal bumps and spines. Unique to echinoderms is the **water vascular system**, a network of hydraulic canals branching into extensions called **tube feet** that function in locomotion and feeding **(Figure 33.44)**. Sexual reproduction of echinoderms usually involves separate male and female individuals that release their gametes into the water.

Echinoderms descended from bilaterally symmetrical ancestors, yet on first inspection most species seem to have a radially symmetrical form. The internal and external parts of most adult echinoderms radiate from the center, often as five spokes. However, echinoderm larvae have bilateral symmetry. Furthermore, the symmetry of adult echinoderms is not truly radial. For example, the opening (madreporite) of a sea star's water vascular system is not central but shifted to one side.

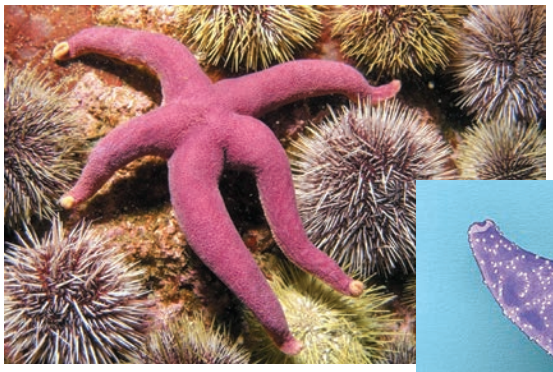
Living echinoderms are divided into five clades.

Asteroidea: Sea Stars and Sea Daisies

Sea stars have arms radiating from a central disk; the undersurfaces of the arms bear tube feet. By a combination of muscular and chemical actions, the tube feet can attach to or detach from a substrate. The sea star adheres firmly to rocks or creeps along slowly as its tube feet extend, grip, release, extend, and grip again. Although the base of the tube foot has a flattened disk that resembles a suction cup, the gripping action results from adhesive chemicals, not suction (see Figure 33.44).

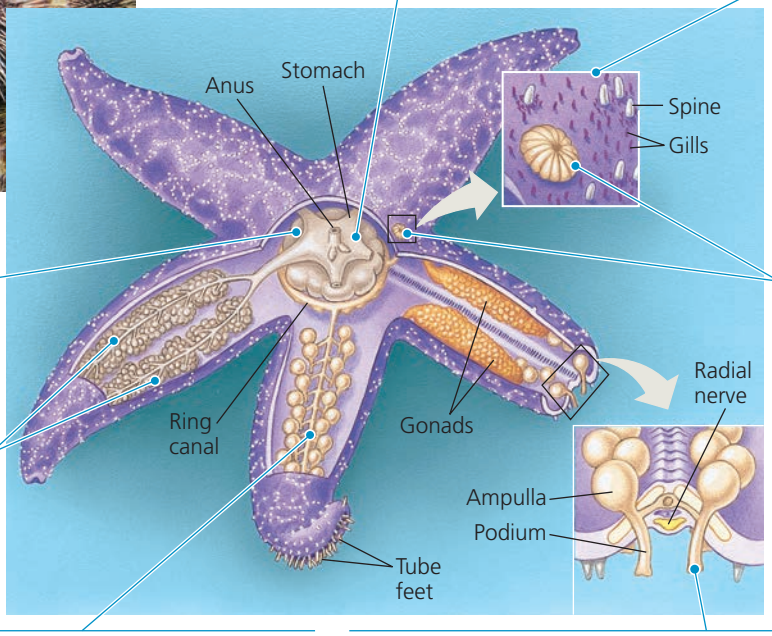
Sea stars also use their tube feet to grasp prey, such as clams and oysters. The arms of the sea star embrace the closed bivalve, clinging tightly with their tube feet. The sea star then turns part of its stomach inside out, everting it through its mouth and into the narrow opening between the halves of the bivalve's shell. Next, the digestive system of the sea star secretes juices that begin digesting the mollusc within its own shell. The sea star then brings its stomach back inside its body, where digestion of the mollusc's (now liquefied) body is completed. The ability to begin the digestive process outside of its body allows a sea star to consume bivalves and other prey species that are much larger than its mouth.

Sea stars and some other echinoderms have considerable powers of regeneration. Sea stars can regrow lost arms, and



A short digestive tract runs from the mouth on the bottom of the central disk to the anus on top of the disk.

The surface of a sea star is covered by spines that help defend against predators, as well as by small gills that provide gas exchange.



Central disk. The central disk has a nerve ring and nerve cords radiating from the ring into the arms.

Madreporite. Water can flow in or out of the water vascular system into the surrounding water through the madreporite.

Digestive glands secrete digestive juices and aid in the absorption and storage of nutrients.

Radial canal. The water vascular system consists of a ring canal in the central disk and five radial canals, each running in a groove down the entire length of an arm. Branching from each radial canal are hundreds of hollow, muscular tube feet filled with fluid.

Each tube foot consists of a bulb-like ampulla and a podium (foot portion). When the ampulla squeezes, water is forced into the podium, which expands and contacts the substrate. Adhesive chemicals are then secreted from the base of the podium, attaching it to the substrate. To detach the tube foot, de-adhesive chemicals are secreted and muscles in the podium contract, forcing water back into the ampulla and shortening the podium. As it moves, a sea star leaves an observable "footprint" of adhesive material on the substrate.

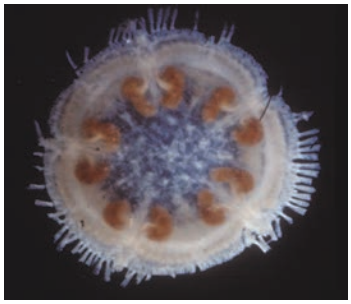
Figure 33.44 Anatomy of a sea star, an echinoderm (top view). The photograph shows a sea star surrounded by sea urchins, which are members of the echinoderm clade Echinoidea.

Video: Echinoderm Tube Feet

members of one genus can even regrow an entire body from a single arm if part of the central disk remains attached.

The clade Asteroidea, to which sea stars belong, also includes a small group of armless species, the *sea daisies*. Only three species of sea daisies are known, all of which live on submerged wood. A sea daisy's body is typically disk-shaped; it has a five-sided organization and measures less than a centimeter in diameter **(Figure 33.45)**. The edge of the body is ringed with small spines. Sea daisies absorb nutrients through a membrane that surrounds their body.

Figure 33.45 A sea daisy (clade Asteroidea).



Ophiuroidea: Brittle Stars

Brittle stars have a distinct central disk and long, flexible arms **(Figure 33.46)**. They move primarily by lashing their arms in serpentine movements. The base of a brittle star tube foot lacks the flattened disk found in sea stars but does secrete adhesive chemicals. Hence, like sea stars and other

Figure 33.46 A brittle star (clade Ophiuroidea).

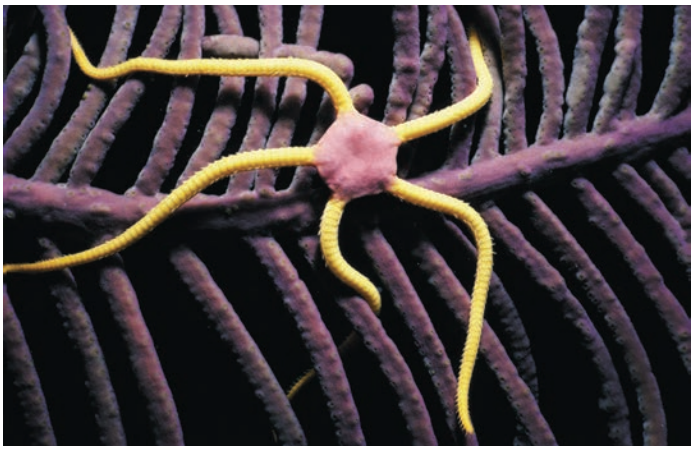
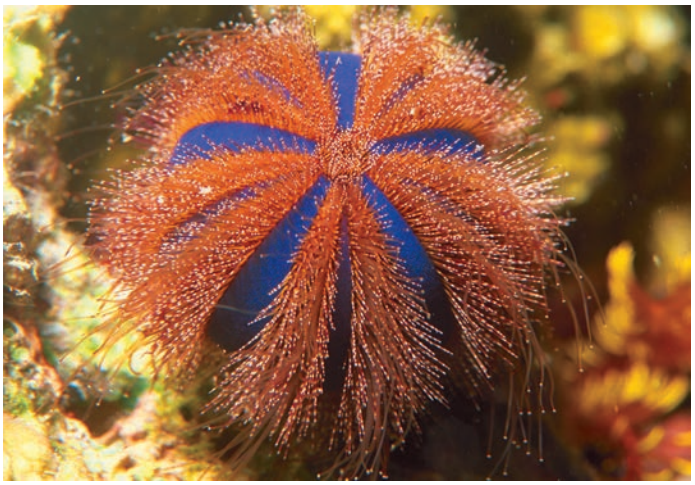


Figure 33.47 **A sea urchin (clade Echinoidea).**



echinoderms, brittle stars can use their tube feet to grip substrates. Some species are suspension feeders; others are predators or scavengers.

Echinoidea: Sea Urchins and Sand Dollars

Sea urchins and sand dollars have no arms, but they do have five radially arranged groups of tube feet that function in slow movement. Sea urchins also have muscles that pivot their long spines, which aid in locomotion as well as protection **(Figure 33.47)**. A sea urchin's mouth, located on its underside, is ringed by highly complex, jaw-like structures that are well adapted to eating seaweed. Sea urchins are roughly spherical, whereas sand dollars are flat disks.

Crinoidea: Sea Lilies and Feather Stars

Sea lilies live attached to the substrate by a stalk; feather stars crawl about by using their long, flexible arms. Both use their arms in suspension feeding. The arms encircle the mouth, which is directed upward, away from the substrate **(Figure 33.48)**. Crinoidea is an ancient group whose

Figure 33.48 **A feather star (clade Crinoidea).**

Figure 33.49 **A sea cucumber (clade Holothuroidea).**

morphology has changed little over the course of evolution; fossilized sea lilies some 500 million years old are extremely similar to present-day members of the clade.

Holothuroidea: Sea Cucumbers

On casual inspection, sea cucumbers do not look much like other echinoderms. They lack spines, and their endoskeleton is much reduced. They are also elongated in their oral-aboral axis, giving them the shape for which they are named and further disguising their relationship to sea stars and sea urchins **(Figure 33.49)**. Closer examination, however, reveals that sea cucumbers have five radially arranged sections of tube feet, as in other echinoderms. Some of the tube feet around the mouth are developed as feeding tentacles.

Chordates

Phylum Chordata consists of two basal groups of invertebrates, the lancelets and the tunicates, as well as the vertebrates. Chordates are bilaterally symmetrical coelomates with segmented bodies. The close relationship between echinoderms and chordates does not mean that one phylum evolved from the other. In fact, echinoderms and chordates have evolved independently of one another for over 500 million years. We will trace the phylogeny of chordates in Chapter 34, focusing on the history of vertebrates.

CONCEPT CHECK 33.5

1. How does a sea star feed on prey larger than its mouth?
2. **WHAT IF? ➤** The insect *Drosophila melanogaster* and the nematode *Caenorhabditis elegans* are prominent model organisms. Are these species the most appropriate invertebrates for making inferences about humans and other vertebrates? Explain.
3. **MAKE CONNECTIONS ➤** Describe how the features and diversity of echinoderms illustrate the unity of life, the diversity of life, and the match between organisms and their environments (see Concept 21.2).

For suggested answers, see Appendix A.

33 Chapter Review

Go to **MasteringBiology**™ for Videos, Animations, Vocab Self-Quiz, Practice Tests, and more in the Study Area.

SUMMARY OF KEY CONCEPTS

This table recaps the animal groups in this chapter.

VOCAB SELF-QUIZ
goo.gl/Rn5Uax

Key Concept	Phylum	Description
CONCEPT 33.1 **Sponges are basal animals that lack tissues** *(pp. 742–743)* ? *Lacking tissues and organs, how do sponges accomplish tasks such as gas exchange, nutrient transport, and waste disposal?*	Porifera (sponges)	 Lack tissues; have choanocytes (collar cells—flagellated cells that ingest bacteria and tiny food particles)
CONCEPT 33.2 **Cnidarians are an ancient phylum of eumetazoans** *(pp. 743–745)* ? *Describe the cnidarian body plan and its two major variations.*	Cnidaria (hydras, jellies, sea anemones, corals)	 Unique stinging structures (nematocysts) housed in specialized cells (cnidocytes); diploblastic; radially symmetrical; gastrovascular cavity (digestive compartment with a single opening)
CONCEPT 33.3 **Lophotrochozoans, a clade identified by molecular data, have the widest range of animal body forms** *(pp. 746–757)* ? *Is the lophotrochozoan clade united by unique morphological features shared by all of its members? Explain.*	Platyhelminthes (flatworms)	 Dorsoventrally flattened acoelomates; gastrovascular cavity or no digestive tract
	Syndermata (rotifers and acanthocepha-lans)	 Pseudocoelomates. Rotifers have alimentary canal (digestive tube with mouth and anus) and jaws (trophi); acanthocephalans are parasites of vertebrates
	Lophophorates: Ectoprocta, Brachiopoda	 Coelomates with lophophores (feeding structures bearing ciliated tentacles)
	Mollusca (clams, snails, squids)	 Coelomates with three main body parts (muscular foot, visceral mass, mantle); coelom reduced; most have hard shell made of calcium carbonate
	Annelida (segmented worms)	 Coelomates with segmented body wall and internal organs (except di-gestive tract, which is unsegmented)
CONCEPT 33.4 **Ecdysozoans are the most species-rich animal group** *(pp. 757–765)* ? *Describe some ecological roles of nematodes and arthropods.*	Nematoda (roundworms)	 Cylindrical pseudocoelomates with tapered ends; no circulatory system; undergo ecdysis
	Arthropoda (spi-ders, centipedes, crustaceans, and insects)	 Coelomates with segmented body, jointed appendages, and exoskeleton made of protein and chitin
CONCEPT 33.5 **Echinoderms and chordates are deuterostomes** *(pp. 765–767)* ? *You've read that echinoderms and chordates are closely related and have evolved independently for over 500 million years. Explain how both of these statements can be correct.*	Echinodermata (sea stars, sea urchins)	 Coelomates with bilaterally symmetrical larvae and five-part body organization as adults; unique water vascular system; endoskeleton
	Chordata (lancelets, tunicates, vertebrates)	 Coelomates with notochord; dorsal, hollow nerve cord; pharyngeal slits; post-anal tail (see Chapter 34)

Metazoa
Eumetazoa
Bilateria
Lophotrochozoa
Ecdysozoa
Deuterostomia

TEST YOUR UNDERSTANDING

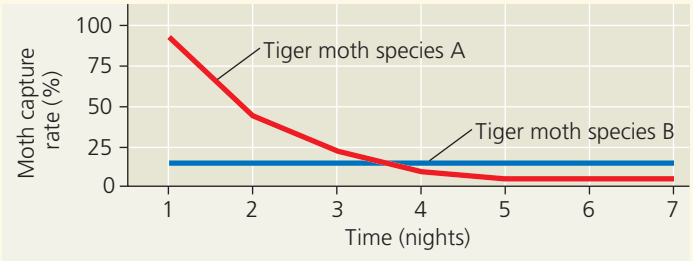
Multiple-choice Self-Quiz questions 1–6 can be found in the Study Area in MasteringBiology.

7. **EVOLUTION CONNECTION • INTERPRET THE DATA** Draw a phylogenetic tree of Bilateria that includes the ten phyla of bilaterians discussed in detail in this chapter. Label each branch that leads to a phylum with a C, P, or A, depending on whether members of the phylum are coelomates (C), pseudocoelomates (P), or acoelomates (A). Use your labeled tree to answer the following questions:

PRACTICE TEST
goo.gl/iAsVgL

   (a) For each of the three major clades of bilaterians, what (if anything) can be inferred about whether the common ancestor of the clade had a true coelom?
   (b) To what extent has the presence of a true coelom in animals changed over the course of evolution?

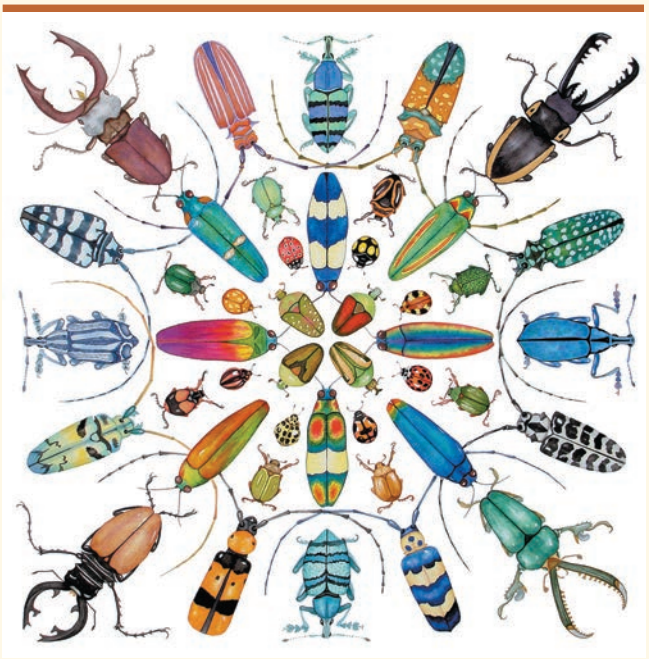
8. **SCIENTIFIC INQUIRY** Bats emit ultrasonic sounds and then use the returning echoes of those sounds to locate and capture flying insects, such as moths, in the dark. In response to bat attacks, some tiger moths make ultrasonic clicks of their own. Researchers hypothesize that tiger moth clicks likely either (1) jam the bat's sonar or (2) warn the bat about the moth's toxic chemical defenses. The graph below shows two patterns observed in studies of moth capture rates over time.



Bats in these experiments were "naive," meaning that prior to the study the bats had not previously hunted tiger moths. Indicate whether the results support hypothesis (1), hypothesis (2), or both. Explain why the researchers used naive bats in this study.

9. **WRITE ABOUT A THEME: ORGANIZATION** Write a short essay (100–150 words) that explains how the respiratory organs in different invertebrate groups have diversified based on their habitats and complexity of body structures.

10. **SYNTHESIZE YOUR KNOWLEDGE**



Collectively, do these beetles and all other invertebrate species combined form a monophyletic group? Explain your answer and provide an overview of the evolutionary history of invertebrate life.

For selected answers, see Appendix A.

For additional practice questions, check out the **Dynamic Study Modules** in MasteringBiology. You can use them to study on your smartphone, tablet, or computer anytime, anywhere!

34 Vertebrates

Figure 34.1 What is the relationship between this ancient organism and humans?

KEY CONCEPTS

34.1 Chordates have a notochord and a dorsal, hollow nerve cord

34.2 Vertebrates are chordates that have a backbone

34.3 Gnathostomes are vertebrates that have jaws

34.4 Tetrapods are gnathostomes that have limbs

34.5 Amniotes are tetrapods that have a terrestrially adapted egg

34.6 Mammals are amniotes that have hair and produce milk

34.7 Humans are mammals that have a large brain and bipedal locomotion

Half a Billion Years of Backbones

Early in the Cambrian period, some 530 million years ago, an immense variety of invertebrate animals inhabited Earth's oceans. Predators used sharp claws and mandibles to capture and break apart their prey. Many animals had protective spikes or armor as well as modified mouthparts that enabled their bearers to filter food from the water.

Amidst this bustle, it would have been easy to overlook certain slender, 3-cm-long creatures gliding through the water: members of the species *Myllokunmingia fengjiaoa* **(Figure 34.1)**. Although lacking armor and appendages, this ancient species was closely related to one of the most successful groups of animals ever to swim, walk, slither, or fly: the **vertebrates**, which derive their name from vertebrae, the series of bones that make up the vertebral column, or backbone.

For more than 150 million years, vertebrates were restricted to the oceans, but about 365 million years ago, the evolution of limbs in one lineage of vertebrates set the stage for these vertebrates to colonize land. Over time, as the descendants of these early colonists adapted to life on land, they gave rise to the three groups of terrestrial vertebrates alive today: the amphibians, the reptiles (including birds), and the mammals.

When you see this blue icon, log in to **MasteringBiology** and go to the Study Area for digital resources.

Get Ready for This Chapter

There are more than 57,000 species of vertebrates, a relatively small number compared to, say, the 1 million insect species on Earth. But what vertebrates may lack in number of species, they make up for in *disparity*, varying enormously in characteristics such as body mass. Vertebrates include the heaviest animals ever to walk on land, plant-eating dinosaurs that were as massive as 40,000 kg (more than 13 pickup trucks). The biggest animal ever to exist on Earth is also a vertebrate—the blue whale, which can exceed 100,000 kg. On the other end of the spectrum, the fish *Schindleria brevipinguis* is just 8.4 mm long and has a mass roughly 100 billion times smaller than that of a blue whale.

In this chapter, you will learn about current hypotheses regarding the origins of vertebrates from invertebrate ancestors. We will track the evolution of the vertebrate body plan, from a notochord to a head to a mineralized skeleton. We'll also explore the major groups of vertebrates (both living and extinct), as well as the evolutionary history of our own species—*Homo sapiens*.

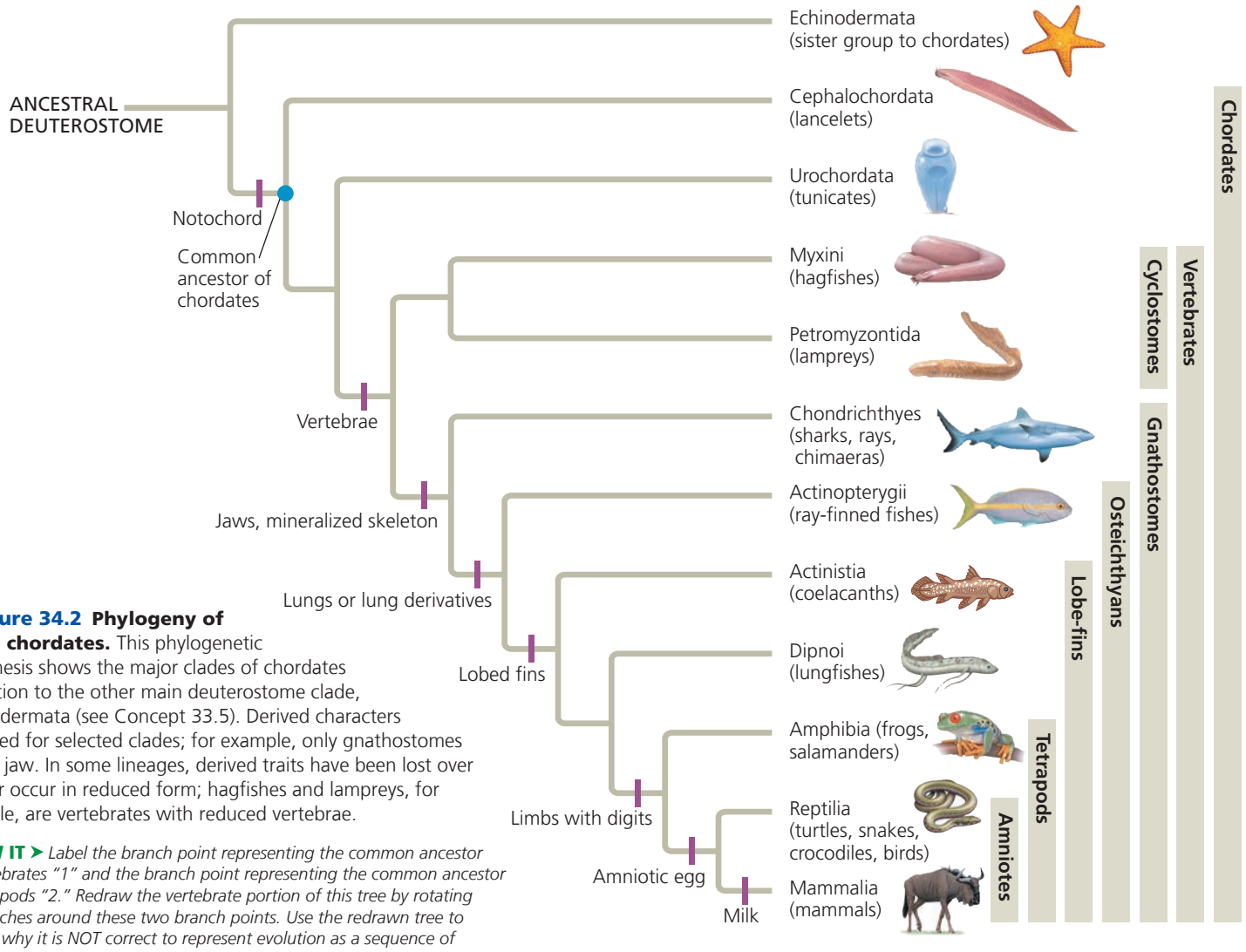
CONCEPT 34.1

Chordates have a notochord and a dorsal, hollow nerve cord

Vertebrates are members of the phylum Chordata, the chordates. **Chordates** are bilaterian (bilaterally symmetrical) animals, and within Bilateria, they belong to the clade of animals known as Deuterostomia (see Figure 32.11). As shown in **Figure 34.2**, there are two groups of invertebrate deuterostomes that are more closely related to vertebrates than they are to other invertebrates: the cephalochordates and the urochordates. Thus, along with the vertebrates, these two invertebrate groups are classified within the chordates.

Derived Characters of Chordates

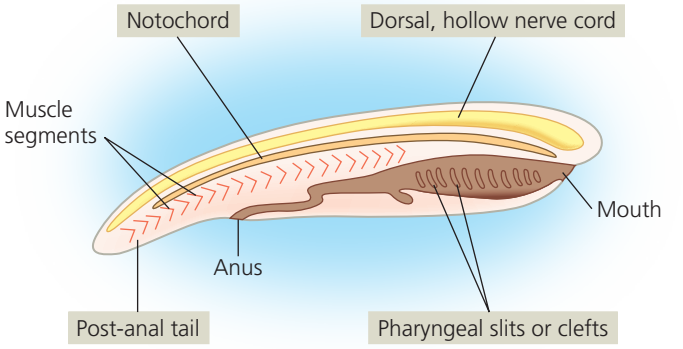
All chordates share a set of derived characters, though many species possess some of these traits only during embryonic



➤ Figure 34.2 Phylogeny of living chordates. This phylogenetic hypothesis shows the major clades of chordates in relation to the other main deuterostome clade, Echinodermata (see Concept 33.5). Derived characters are listed for selected clades; for example, only gnathostomes have a jaw. In some lineages, derived traits have been lost over time or occur in reduced form; hagfishes and lampreys, for example, are vertebrates with reduced vertebrae.

DRAW IT ➤ *Label the branch point representing the common ancestor of vertebrates "1" and the branch point representing the common ancestor of tetrapods "2." Redraw the vertebrate portion of this tree by rotating its branches around these two branch points. Use the redrawn tree to explain why it is NOT correct to represent evolution as a sequence of events "leading to" humans and other mammals.*

Figure 34.3 Chordate characteristics. All chordates possess the four highlighted structural trademarks at some point during their development.



development. **Figure 34.3** illustrates four key characters of chordates: a notochord; a dorsal, hollow nerve cord; pharyngeal slits or clefts; and a muscular, post-anal tail.

Notochord

Chordates are named for a skeletal structure, the notochord, present in all chordate embryos as well as in some adult chordates. The **notochord** is a longitudinal, flexible rod located between the digestive tube and the nerve cord. It is composed of large, fluid-filled cells encased in fairly stiff, fibrous tissue. The notochord provides skeletal support throughout most of the length of a chordate, and in larvae or adults that retain it, it also provides a firm but flexible structure against which muscles can work during swimming. In most vertebrates, a more complex, jointed skeleton develops around the ancestral notochord, and the adult retains only remnants of the embryonic notochord. In humans, for example, the notochord is reduced and forms part of the gelatinous disks sandwiched between the vertebrae.

Dorsal, Hollow Nerve Cord

The nerve cord of a chordate embryo develops from a plate of ectoderm that rolls into a neural tube located dorsal to the notochord. The resulting dorsal, hollow nerve cord is unique to chordates. Other animal phyla have solid nerve cords, and in most cases they are ventrally located. The nerve cord of a chordate embryo develops into the central nervous system: the brain and spinal cord.

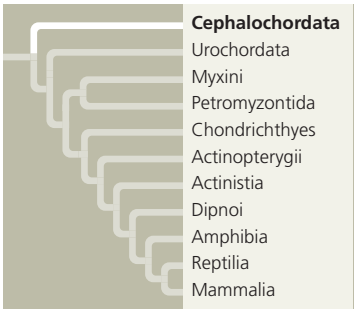
Pharyngeal Slits or Clefts

The digestive tube of chordates extends from the mouth to the anus. The region just posterior to the mouth is the pharynx. In all chordate embryos, a series of arches separated by grooves forms along the outer surface of the pharynx. In most chordates, these grooves (known as **pharyngeal clefts**) develop into slits that open into the pharynx. These **pharyngeal slits** allow water entering the mouth to exit the body without passing through the entire digestive tract. Pharyngeal slits function as suspension-feeding devices in many invertebrate chordates. In vertebrates (with the exception of vertebrates with limbs, the *tetrapods*), these slits and the pharyngeal arches that support them have been modified for gas exchange and are called gills. In tetrapods, the pharyngeal clefts do not develop into slits. Instead, the pharyngeal arches that surround the clefts develop into parts of the ear and other structures in the head and neck.

Muscular, Post-Anal Tail

Chordates have a tail that extends posterior to the anus, although in many species it is greatly reduced during embryonic development. In contrast, most nonchordates have a digestive tract that extends nearly the whole length of the body. The chordate tail contains skeletal elements and muscles, and it helps propel many aquatic species in the water.

Lancelets

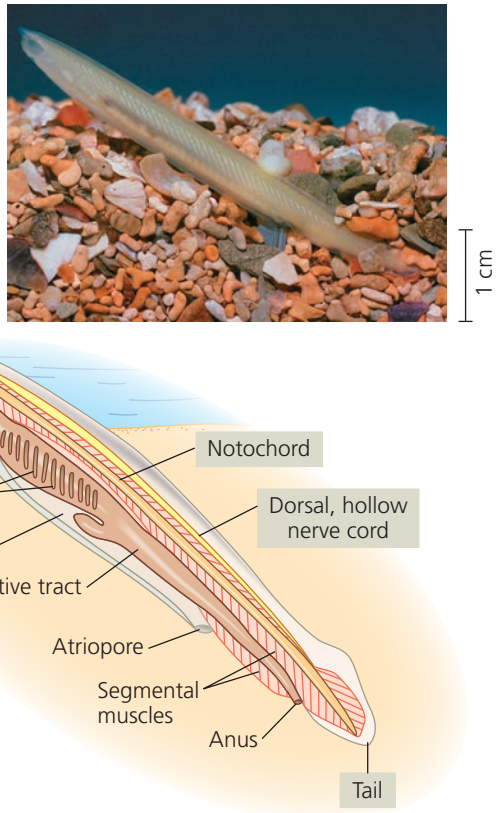


The most basal (earliest-diverging) group of living chordates are animals called **lancelets** (Cephalochordata), which get their name from their bladelike shape **(Figure 34.4)**. As larvae, lancelets develop a notochord; a dorsal, hollow nerve cord; numerous pharyngeal slits; and a post-anal tail. The larvae feed on plankton in the water column, alternating between upward swimming and passive sinking. As the larvae sink, they trap plankton and other suspended particles in their pharynx.

Adult lancelets can reach 6 cm in length. They retain key chordate traits, closely resembling the idealized chordate shown in Figure 34.3. Following metamorphosis, an adult lancelet swims down to the seafloor and wriggles backward into the sand, leaving only its anterior end exposed. Cilia draw seawater into the lancelet's mouth. A net of mucus secreted across the pharyngeal slits removes tiny food particles as the water passes through the slits, and the trapped food enters the intestine. The pharynx and pharyngeal slits play a minor role in gas exchange, which occurs mainly across the external body surface.

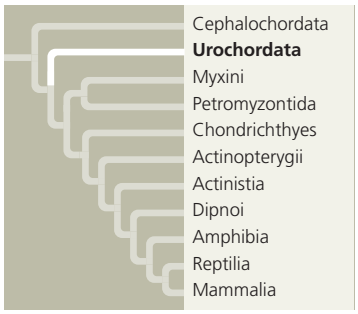
A lancelet frequently leaves its burrow to swim to a new location. Though feeble swimmers, these invertebrate chordates display, in a simple form, the swimming mechanism of fishes. Coordinated contraction of muscles arranged like rows of chevrons (>>>>) along the sides of the notochord flexes the notochord, producing side-to-side undulations that thrust the body forward. This serial arrangement of muscles is evidence of the lancelet's segmentation. The muscle segments develop from blocks of mesoderm called *somites*, which are found along each side of the notochord in all chordate embryos.

Figure 34.4 The lancelet *Branchiostoma*, a cephalochordate. This small invertebrate displays all four main chordate characters. Water enters the mouth and passes through the pharyngeal slits into the atrium, a chamber that vents to the outside via the atriopore; large particles are blocked from entering the mouth by tentacle-like cirri. The serially arranged segmental muscles produce the lancelet's wavelike swimming movements.



Globally, lancelets are rare, but in a few areas (such as Tampa Bay, on the Florida coast), they may reach densities of more than 5,000 individuals per square meter.

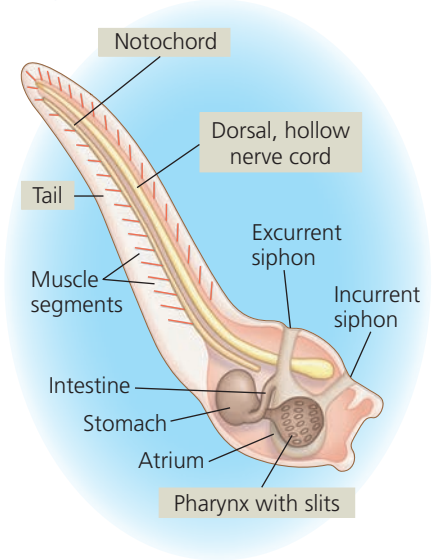
Tunicates



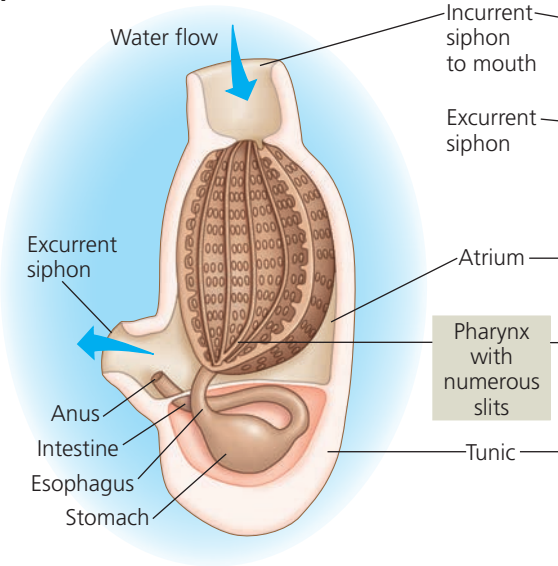
Recent molecular studies indicate that the **tunicates** (Urochordata) are more closely related to other chordates than are lancelets. The chordate characters of tunicates are most apparent during their larval stage, which may be as brief as a few minutes **(Figure 34.5a)**. In many species, the larva uses its tail muscles and notochord to swim through water in search of a suitable substrate on which it can settle, guided by cues it receives from light- and gravity-sensitive cells.

Once a tunicate has settled on a substrate, it undergoes a radical metamorphosis in which many of its chordate characters disappear. Its tail and notochord are resorbed; its nervous system degenerates; and its remaining organs rotate 90°. As an adult, a tunicate draws in water through an incurrent siphon; the water then passes through the pharyngeal slits into a chamber called the atrium and exits through an excurrent siphon **(Figure 34.5b** and **c)**. Food particles are filtered from the water by a mucous net and transported by cilia to the esophagus. The anus empties into the excurrent siphon. Some tunicate species shoot a jet of water through their excurrent siphon when attacked, earning them the informal name of "sea squirts."

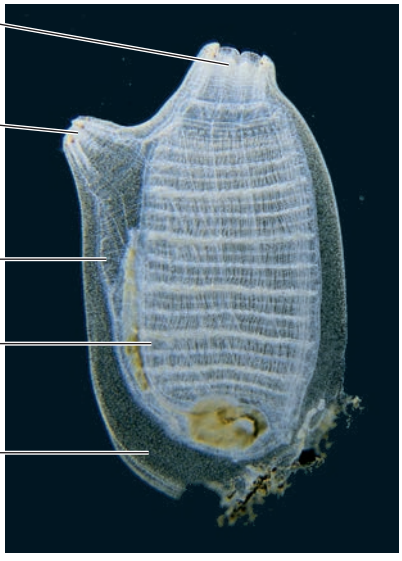
Figure 34.5 A tunicate, a urochordate.



(a) A tunicate larva is a free-swimming but nonfeeding "tadpole" in which all four main characters of chordates are evident.



(b) In the adult, prominent pharyngeal slits function in suspension feeding, but other chordate characters are not obvious.



(c) An adult tunicate, or sea squirt, is a sessile animal (photo is approximately life-sized).

The loss of chordate characters in the adult stage of tunicates appears to have occurred after the tunicate lineage branched off from other chordates. Even the tunicate larva appears to be highly derived. For example, tunicates have nine *Hox* genes, whereas all other chordates studied to date—including the early-diverging lancelets—share a set of 13 *Hox* genes. The apparent loss of four *Hox* genes indicates that the chordate body plan of a tunicate larva is built using a different set of genetic controls than other chordates.

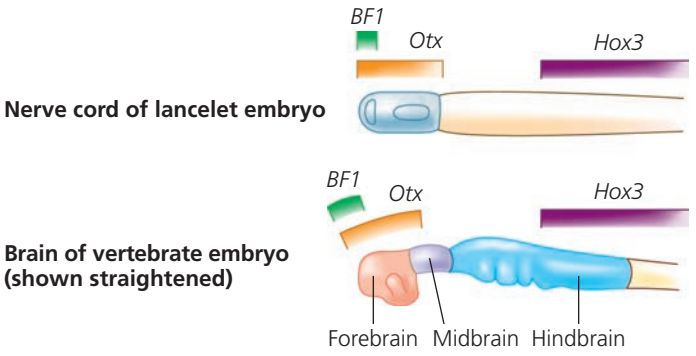
Early Chordate Evolution

Although lancelets and tunicates are relatively obscure animals, they occupy key positions in the history of life and can provide clues about the evolutionary origin of vertebrates. For example, as you have read, lancelets display key chordate characters as adults, and their lineage branches from the base of the chordate phylogenetic tree. These findings suggest that the ancestral chordate may have looked something like a lancelet—that is, it had an anterior end with a mouth; a notochord; a dorsal, hollow nerve cord; pharyngeal slits; and a post-anal tail.

Research on lancelets has also revealed important clues about the evolution of the chordate brain. Rather than a full-fledged brain, lancelets have only a slightly swollen tip on the anterior end of their dorsal nerve cord **(Figure 34.6)**. But the same *Hox* genes that organize major regions of the forebrain, midbrain, and hindbrain of vertebrates express themselves in a corresponding pattern in this small cluster of cells in the lancelet's nerve cord. This suggests that the vertebrate brain is an elaboration of an ancestral structure similar to the lancelet's simple nerve cord tip.

As for tunicates, several of their genomes have been completely sequenced and can be used to identify genes likely to have been present in early chordates. Researchers have suggested that ancestral chordates had genes associated with vertebrate organs such as the heart and thyroid gland. These genes are found in tunicates and vertebrates but are absent from nonchordate invertebrates. In another example, a 2015 study found that tunicates (but not lancelets) have embryonic cells that have some of the characteristics of the *neural crest*, a derived trait found in all vertebrates (see Figure 34.7). This suggests that embryonic cells similar to those in tunicates may represent an intermediate cell population from which the vertebrate neural crest evolved.

Figure 34.6 Expression of developmental genes in lancelets and vertebrates. *Hox* genes (including *BF1*, *Otx*, and *Hox3*) control the development of major regions of the vertebrate brain. These genes are expressed in the same anterior-to-posterior order in lancelets and vertebrates. Each colored bar is positioned above the portion of the brain whose development that gene controls.



MAKE CONNECTIONS ➤ *What do these expression patterns and those in Figure 20.19 indicate about* Hox *genes and their evolution?*

CONCEPT CHECK 34.1

1. Identify four derived characters that all chordates have at some point during their life.
2. You are a chordate, yet you lack most of the main derived characters of chordates. Explain.
3. **VISUAL SKILLS** ➤ Based on the phylogenetic tree diagram in Figure 34.2, predict which vertebrate groups should have lungs or lung derivatives. Explain.

For suggested answers, see Appendix A.

CONCEPT 34.2

Vertebrates are chordates that have a backbone

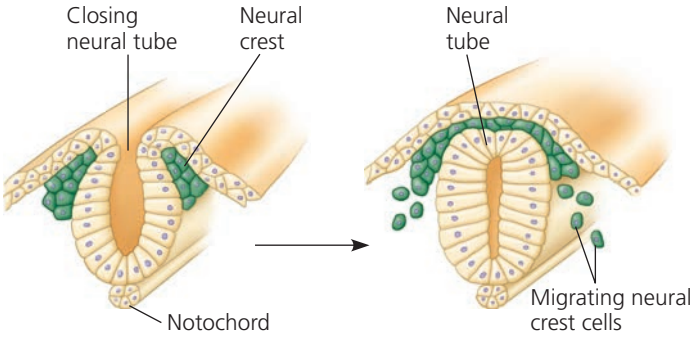
During the Cambrian period, half a billion years ago, a lineage of chordates gave rise to vertebrates. With a skeletal system and a more complex nervous system than that of their ancestors, vertebrates became more efficient at two essential tasks: capturing food and avoiding being eaten.

Derived Characters of Vertebrates

Living vertebrates share a set of derived characters that distinguish them from other chordates. For example, as a result of gene duplication, vertebrates possess two or more sets of *Hox* genes (lancelets and tunicates have only one). Other important families of genes that produce transcription factors and signaling molecules are also duplicated in vertebrates. The resulting additional genetic complexity may be associated with innovations in the vertebrate nervous system and skeleton, including the development of a skull and a backbone composed of vertebrae. In some vertebrates, the vertebrae are little more than small prongs of cartilage arrayed dorsally along the notochord. In the majority of vertebrates, however, the vertebrae enclose the spinal cord and have taken over the mechanical roles of the notochord.

Another feature unique to vertebrates is the **neural crest**, a collection of cells that appears along the edges of the closing neural tube of an embryo **(Figure 34.7)**. Neural crest cells disperse throughout the embryo, where they give rise to a variety of structures, including teeth, some of the bones and cartilage of the skull, several types of neurons, and the sensory capsules in which eyes and other sense organs develop.

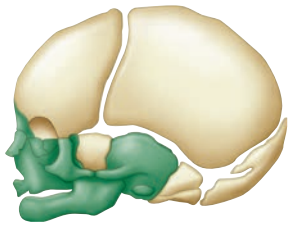
Figure 34.7 The neural crest, embryonic source of many unique vertebrate traits.



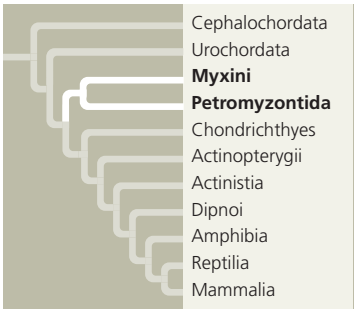
(a) The neural crest consists of bilateral bands of cells near the margins of the embryonic folds that form the neural tube.

(b) Neural crest cells migrate to distant sites in the embryo.

(c) The migrating neural crest cells give rise to some of the anatomical structures unique to vertebrates, including some of the bones and cartilage of the skull. (A fetal human skull is depicted here.)



Hagfishes and Lampreys



The **hagfishes** (Myxini) and the **lampreys** (Petromyzontida) are the only lineages of living vertebrates whose members lack jaws. Unlike most vertebrates, lampreys and hagfishes also do not have a backbone. Even so, lampreys are classified as vertebrates because they have rudimentary vertebrae (composed of cartilage, not bone). The hagfishes, in contrast, traditionally were thought to lack vertebrae altogether; hence, they were classified as invertebrate chordates closely related to vertebrates.

In the past few years, however, this interpretation has changed. Recent research has shown that hagfishes, like lampreys, have rudimentary vertebrae. In addition, a series of molecular phylogenetic studies have supported the hypothesis that hagfishes are vertebrates. Molecular analyses also indicate that hagfishes and lampreys are sister groups, as shown in Figure 34.2. Together, the hagfishes and lampreys form a clade of living jawless vertebrates, the **cyclostomes**. (All other vertebrates have jaws and make up a much larger clade, the gnathostomes, which we will discuss in Concept 34.3.)

Hagfishes

The hagfishes are jawless vertebrates that have highly reduced vertebrae and a skull that is made of cartilage. They swim in a snakelike fashion by using their segmental muscles to exert force against their notochord, which they retain in adulthood as a strong, flexible rod of cartilage. Hagfishes have a small brain, eyes, ears, and a nasal opening that connects with the pharynx. Their mouths contain tooth-like formations made of the protein keratin.

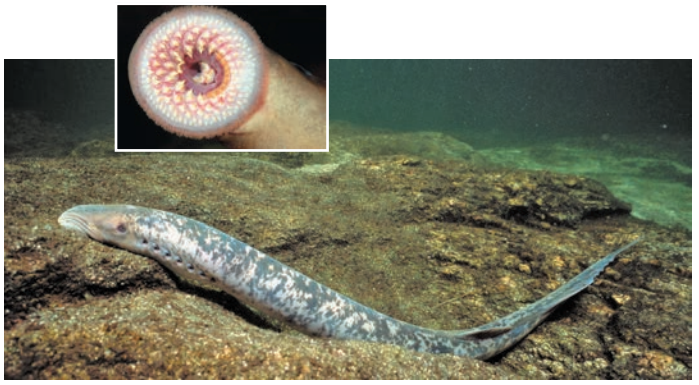
Figure 34.8 A hagfish.

All of the 30 living species of hagfishes are marine. Measuring up to 60 cm in length, most are bottom-dwelling scavengers **(Figure 34.8)** that feed on worms and sick or dead fish. Rows of slime glands on a hagfish's flanks secrete a substance that absorbs water, forming a slime that may repel other scavengers when a hagfish is feeding. When attacked by a predator, a hagfish can produce several liters of slime in less than a minute. The slime coats the gills of the attacking fish, sending it into retreat or even suffocating it. Biologists and engineers are investigating the properties of hagfish slime as a model for developing a space-filling gel that could be used, for instance, to stop bleeding during surgery.

Lampreys

The second group of living jawless vertebrates, the lampreys, consists of about 38 species inhabiting various marine and freshwater environments **(Figure 34.9)**. Some are parasites

Figure 34.9 A sea lamprey. Parasitic lampreys use their mouth (inset) and tongue to bore a hole in the side of a fish. The lamprey then ingests the blood and other tissues of its host.



that feed by clamping their round, jawless mouth onto the flank of a live fish, their "host." Parasitic lampreys use their rasping mouth and tongue to penetrate the skin of the fish and ingest the fish's blood and other tissues.

As larvae, lampreys live in freshwater streams. The larva is a suspension feeder that resembles a lancelet and spends much of its time partially buried in sediment. About 20 species of lampreys are not parasitic. These species feed only as larvae; following several years in streams, they mature sexually, reproduce, and die within a few days. In contrast, parasitic species of lampreys migrate to the sea or lakes as they mature into adults. One such parasite, the sea lamprey (*Petromyzon marinus*), has invaded the Great Lakes over the past 170 years and devastated a number of fisheries there.

The skeleton of lampreys is made of cartilage. Unlike the cartilage found in most vertebrates, lamprey cartilage contains no collagen. Instead, it is a stiff matrix of other proteins. The notochord of lampreys persists as the main axial skeleton in the adult, as it does in hagfishes. However, lampreys also have a flexible sheath around their rodlike notochord. Along the length of this sheath, pairs of cartilaginous projections related to vertebrae extend dorsally, partially enclosing the nerve cord.

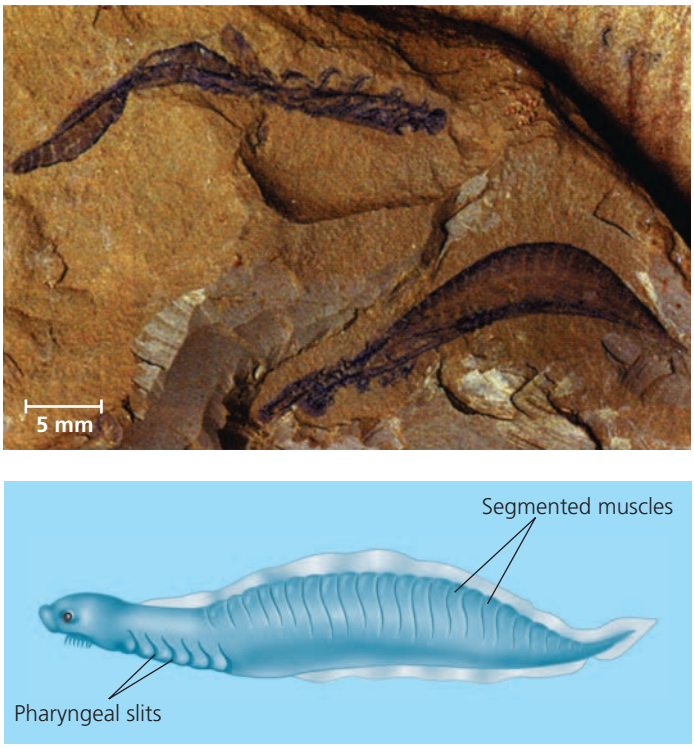
Early Vertebrate Evolution

In the late 1990s, paleontologists working in China discovered a vast collection of fossils of early chordates that appear to straddle the transition to vertebrates. The fossils were formed during the Cambrian explosion 530 million years ago, when many animal groups were undergoing rapid diversification (see Concept 32.2).

The most primitive of the fossils are the 3-cm-long *Haikouella* **(Figure 34.10)**. In many ways, *Haikouella* resembled a lancelet. Its mouth structure indicates that, like lancelets, it probably was a suspension feeder. However, *Haikouella* also had some of the characters of vertebrates. For example, it had a well-formed brain, small eyes, and muscle segments along the body, as do the vertebrate fishes. Unlike the vertebrates, however, *Haikouella* did not have a skull or ear organs, suggesting that these characters emerged with further innovations to the chordate nervous system. (The earliest "ears" were organs for maintaining balance, a function still performed by the ears of humans and other living vertebrates.)

Early signs of a skull can be seen in *Myllokunmingia* (see Figure 34.1). About the same size as *Haikouella, Myllokunmingia* had ear capsules and eye capsules, parts of the skull that surround these organs. Based on these and other characters, *Myllokunmingia* is considered the first chordate to have a head. The origin of a head—consisting of a brain at the anterior end of the dorsal nerve cord, eyes and other sensory organs, and a skull—enabled chordates to coordinate more complex movement and feeding behaviors. Although it had a head, *Myllokunmingia* lacked vertebrae and hence is not classified as a vertebrate.

Figure 34.10 Fossil of an early chordate. Discovered in 1999 in southern China, *Haikouella* had eyes and a brain but lacked a skull, a trait found in vertebrates. The organism's color in the drawing is fanciful.



The earliest fossils of vertebrates date to 500 million years ago and include those of **conodonts**, a group of slender, soft-bodied vertebrates that lacked jaws and whose internal skeleton was composed of cartilage. Conodonts had large eyes, which they may have used in locating prey that were then impaled on a set of barbed hooks at the anterior end of their mouth **(Figure 34.11)**. These hooks were made of

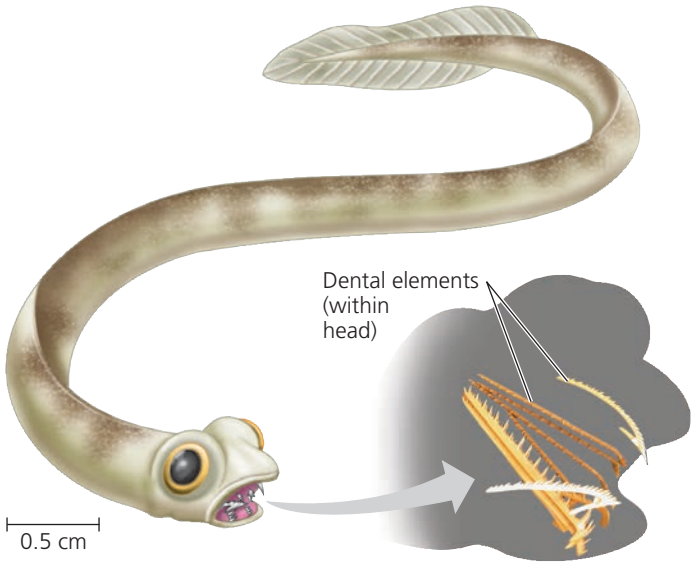


Figure 34.11 A conodont. Conodonts were early jawless vertebrates that lived from 500 million to 200 million years ago. Unlike hagfishes and lampreys, conodonts had mineralized mouthparts, which they used for either predation or scavenging.

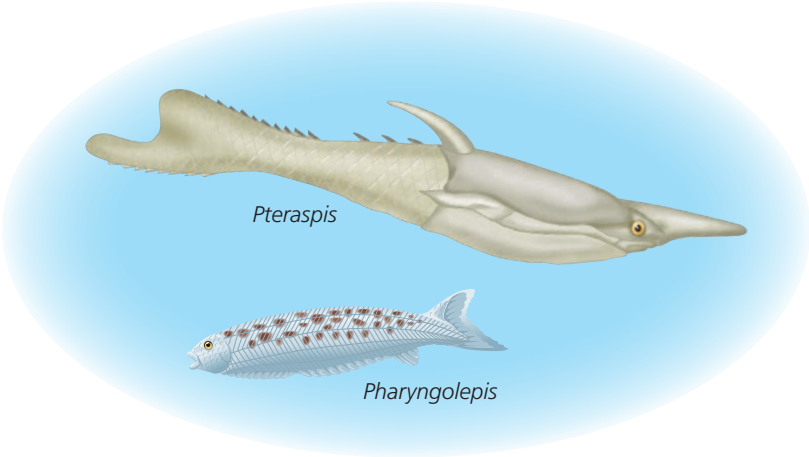
dental tissues that were *mineralized*—hardened by the incorporation of minerals such as calcium. The food was then passed back to the pharynx, where a different set of dental elements sliced and crushed the food.

Conodonts were extremely abundant for 300 million years. Their fossilized dental elements are so plentiful that they have been used for decades by petroleum geologists as guides to the age of rock layers in which they search for oil.

Vertebrates with additional innovations emerged during the Ordovician, Silurian, and Devonian periods (485–359 million years ago). These vertebrates had paired fins and, as in lampreys, an inner ear with two semicircular canals that provided a sense of balance. Like conodonts, these vertebrates lacked jaws, but they had a muscular pharynx, which they may have used to suck in bottom-dwelling organisms or detritus. They were also armored with mineralized bone, which covered varying amounts of their body and may have offered protection from predators **(Figure 34.12)**. There were many species of these jawless, armored swimming vertebrates, but they all became extinct by the end of the Devonian.

Finally, note that the human skeleton is heavily mineralized bone, whereas cartilage plays a fairly minor role. But a bony internal skeleton was a relatively late development in the history of vertebrates. Instead, the vertebrate skeleton evolved initially as a structure made of unmineralized cartilage. Steps toward a bony skeleton began 470 million years ago, with the appearance of mineralized bone on the outer surface of the skull in some jawless vertebrates. Shortly after that time, the internal skeleton began to mineralize, first as calcified cartilage. By 430 million years ago, some vertebrates had a thin layer of bone lining the cartilage of their internal skeleton. The bones of vertebrates underwent even more mineralization in the group we turn to next, the jawed vertebrates.

Figure 34.12 Jawless armored vertebrates. *Pteraspis* and *Pharyngolepis* were two of many genera of jawless vertebrates that emerged during the Ordovician, Silurian, and Devonian periods.



CONCEPT CHECK 34.2

1. How are differences in the anatomy of lampreys and conodonts reflected in each animal's feeding method?
2. **WHAT IF?** ➤ In several different animal lineages, organisms with a head first appeared around 530 million years ago. Does this finding constitute proof that having a head is favored by natural selection? Explain.
3. **WHAT IF?** ➤ Suggest key roles that mineralized bone might have played in early vertebrates.

For suggested answers, see Appendix A.

CONCEPT 34.3

Gnathostomes are vertebrates that have jaws

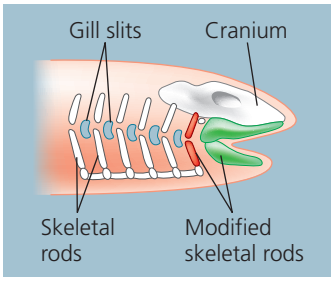
Hagfishes and lampreys are survivors from the early Paleozoic era, when jawless vertebrates were common. Since then, jawless vertebrates have been far outnumbered by the jawed vertebrates, the **gnathostomes**. Living gnathostomes are a diverse group that includes sharks and their relatives, ray-finned fishes, lobe-finned fishes, amphibians, reptiles (including birds), and mammals.

Derived Characters of Gnathostomes

Gnathostomes ("jaw mouth") are named for their jaws, hinged structures that, especially with the help of teeth, enable gnathostomes to grip food items firmly and slice them. According to one hypothesis, gnathostome jaws evolved by modification of the skeletal rods that had previously supported the anterior pharyngeal (gill) slits. **Figure 34.13** shows a stage in this evolutionary process in which several of these skeletal rods have been modified into precursors of jaws (green) and their structural supports (red). The remaining gill slits, no longer required for suspension feeding, remained as the major sites of respiratory gas exchange with the external environment.

Gnathostomes share other derived characters besides jaws. The common ancestors of all gnathostomes underwent an additional duplication of *Hox* genes, such that the single set present in early chordates became four. In fact, the entire genome appears to have duplicated, and together these genetic changes likely enabled the origin of jaws and other novel features in gnathostomes. The gnathostome forebrain is enlarged compared to that of other vertebrates, and it is associated with enhanced senses of smell and vision. Another characteristic of aquatic

Figure 34.13 Possible step in the evolution of jawbones.



gnathostomes is the **lateral line system**, organs that form a row along each side of the body and are sensitive to vibrations in the surrounding water. Precursors of these organs were present in the head shields of some jawless vertebrates.

Fossil Gnathostomes

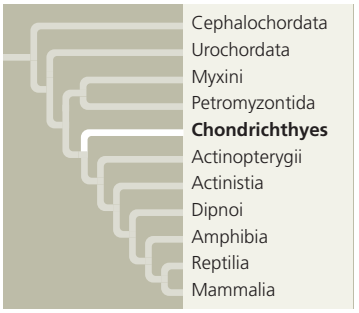
Gnathostomes appeared in the fossil record about 440 million years ago and steadily became more diverse. Their success probably resulted from a combination of anatomical features: Their paired fins and tail (which were also found in jawless vertebrates) allowed them to swim efficiently after prey, and their jaws enabled them to grab prey or simply bite off chunks of flesh. Over time, dorsal, ventral, and anal fins stiffened by bony structures called fin rays also evolved in some early gnathostomes. Fin rays provide thrust and steering control when aquatic vertebrates swim after prey or away from predators. Faster swimming was supported by other adaptations, including a more efficient gas exchange system in the gills.

The earliest gnathostomes include extinct lineages of armored vertebrates known collectively as **placoderms**, which means "plate-skinned." Most placoderms were less than a meter long, though some giants measured more than 10 m **(Figure 34.14)**. Other jawed vertebrates, called **acanthodians**, emerged at roughly the same time and radiated during the Silurian and Devonian periods (444–359 million years ago). Placoderms had disappeared by 359 million years ago, and acanthodians became extinct about 70 million years later.

Overall, a series of recent fossil discoveries have revealed that 440–420 million years ago was a period of tumultuous evolutionary change. Gnathostomes that lived during this period had highly variable forms, and by 420 million years ago, they had diverged into the three lineages of jawed vertebrates that survive today: chondrichthyans, ray-finned fishes, and lobe-fins.

Figure 34.14 Fossil of an early gnathostome. A formidable predator, the placoderm *Dunkleosteus* grew up to 10 m in length. Its jaw structure indicates that *Dunkleosteus* could exert a force of 560 kg/cm^2 (8,000 pounds per square inch) at the tip of its jaws.

Chondrichthyans (Sharks, Rays, and Their Relatives)



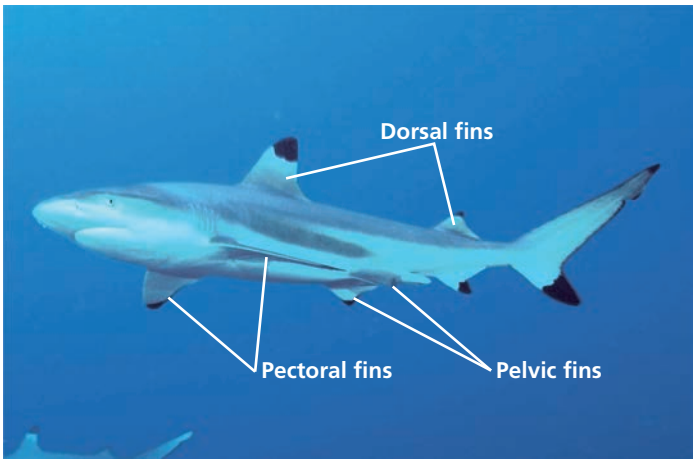
Sharks, rays, and their relatives include some of the biggest and most successful vertebrate predators in the oceans. They belong to the clade Chondrichthyes, which means "cartilage fish." As their name indicates, the **chondrichthyans** have a skeleton composed predominantly of cartilage, though often impregnated with calcium.

When the name Chondrichthyes was first coined in the 1800s, scientists thought that chondrichthyans represented an early stage in the evolution of the vertebrate skeleton and that mineralization had evolved only in more derived lineages (such as "bony fishes"). However, as armored jawless vertebrates demonstrate, the mineralization of the vertebrate skeleton had already begun before the chondrichthyan lineage branched off from other vertebrates. Moreover, bone-like tissues have been found in early chondrichthyans, such as the fin skeleton of a shark that lived in the Carboniferous period. Traces of bone can also be found in living chondrichthyans—in their scales, at the base of their teeth, and, in some sharks, in a thin layer on the surface of their vertebrae. Such findings indicate that the restricted distribution of bone in the chondrichthyan body is a derived condition, emerging after chondrichthyans diverged from other gnathostomes.

There are about 1,000 species of living chondrichthyans. The largest and most diverse group consists of the sharks, rays, and skates (**Figure 34.15a** and **b**). A second group is composed of a few dozen species of ratfishes, also called chimaeras **(Figure 34.15c)**.

Most sharks have a streamlined body and are swift swimmers, but they do not maneuver very well. Powerful movements of the trunk and the tail fin propel them forward. The dorsal fins function mainly as stabilizers, and the paired pectoral (fore) and pelvic (hind) fins are important for maneuvering. Although a shark gains buoyancy by storing a large amount of oil in its huge liver, the animal is still more dense than water, and if it stops swimming it sinks. Continual swimming also ensures that water flows into the shark's mouth and out through the gills, where gas exchange occurs. However, some sharks and many skates and rays spend a good deal of time resting on the seafloor. When resting, they use muscles of their jaws and pharynx to pump water over the gills.

Figure 34.15 Chondrichthyans.



(a) Blacktip reef shark (*Carcharhinus melanopterus*). Sharks are fast swimmers with acute senses. Like all gnathostomes, they have paired pectoral and pelvic fins.

(b) Southern stingray (*Dasyatis americana*). Most rays are bottom-dwellers that feed on molluscs and crustaceans. Some rays cruise in open water and scoop food into their gaping mouths.

(c) Spotted ratfish (*Hydrolagus colliei*). Ratfishes, or chimaeras, typically live at depths greater than 80 m and feed on shrimp, molluscs, and sea urchins. Some species have a venomous spine at the front of their first dorsal fin.

The largest sharks and rays are suspension feeders that consume plankton. Most sharks, however, are carnivores that swallow their prey whole or use their powerful jaws and sharp teeth to tear flesh from animals too large to swallow in one piece. Sharks have several rows of teeth that gradually move to the front of the mouth as old teeth are lost. The digestive tract of many sharks is proportionately shorter than that of many other vertebrates. Within the shark intestine is a *spiral valve*, a corkscrew-shaped ridge that increases surface area and prolongs the passage of food through the digestive tract.

Acute senses are adaptations that go along with the active, carnivorous lifestyle of sharks. Sharks have sharp vision but cannot distinguish colors. The nostrils of sharks, like those of most aquatic vertebrates, open into dead-end cups. They function only for olfaction (smelling), not for breathing. Like some other vertebrates, sharks have a pair of regions in the skin of their head that can detect electric fields generated by the muscle contractions of nearby animals. Like all nonmammalian aquatic vertebrates, sharks have no eardrums, structures that terrestrial vertebrates use to transmit sound waves in air to the auditory organs. Sound reaches a shark through water, and the animal's entire body transmits the sound to the hearing organs of the inner ear.

Shark eggs are fertilized internally. The male has a pair of claspers on its pelvic fins that transfer sperm into the female's reproductive tract. Some species of sharks are **oviparous**; they lay eggs that hatch outside the mother's body. These sharks release their fertilized eggs after encasing them in protective coats. Other species are **ovoviviparous**; they retain the fertilized eggs in the oviduct. Nourished by the egg yolk, the embryos develop into young that are born after hatching within the uterus. A few species are **viviparous**; the young develop within the uterus and obtain nourishment prior to birth by receiving nutrients from the mother's blood through a yolk sac placenta, by absorbing a nutritious fluid produced by the uterus, or by eating other eggs. The reproductive tract of the shark empties along with the excretory system and digestive tract into the **cloaca**, a common chamber that has a single opening to the outside.

Although rays are closely related to sharks, they have adopted a very different lifestyle. Most rays are bottom-dwellers that feed by using their jaws to crush molluscs and crustaceans. They have a flattened shape and use their greatly enlarged pectoral fins like water wings to propel themselves through the water. The tail of many rays is whiplike and, in some species, bears venomous barbs that function in defense.

Chondrichthyans have thrived for over 400 million years. Today, however, they are severely threatened by overfishing. A 2012 report, for example, indicated that shark populations in the Pacific have plummeted by up to 95%, and shark populations that live closest to people have declined the most.

Ray-Finned Fishes and Lobe-Fins

The vast majority of vertebrates belong to the clade of gnathostomes called Osteichthyes. Unlike chondrichthyans, nearly all living **osteichthyans** have an ossified (bony) endoskeleton with a hard matrix of calcium phosphate. Like many other taxonomic names, the name Osteichthyes ("bony fish") was coined long before the advent of phylogenetic systematics. When it was originally defined, the group excluded tetrapods, but we now know that such a taxon would be paraphyletic (see Figure 34.2). Therefore, systematists today include tetrapods along with bony fishes in the clade Osteichthyes. Clearly, the name of the group does not accurately describe all of its members.

This section discusses the aquatic osteichthyans known informally as fishes. Most fishes breathe by drawing water over four or five pairs of gills located in chambers covered by a protective bony flap called the **operculum** **(Figure 34.16)**. Water is drawn into the mouth, through the pharynx, and out between the gills by movement of the operculum and contraction of muscles surrounding the gill chambers.

Most fishes can maintain a buoyancy equal to the surrounding water by filling an air sac known as a **swim bladder**. (If a fish swims to greater depths or toward the surface, where water pressure differs, the fish shuttles gas between its blood and swim bladder, keeping the volume of gas in the bladder constant.) Charles Darwin proposed that the lungs of tetrapods evolved from swim bladders, but strange as it may sound, the opposite seems to be true: Swim bladders arose from lungs. Osteichthyans in many early-branching lineages have lungs, which they use to breathe air as a supplement to gas exchange in their gills. This suggests that lungs arose in early osteichthyans; later, swim bladders evolved from lungs in some lineages.

In nearly all fishes, the skin is covered by flattened, bony scales that differ in structure from the tooth-like scales of sharks. Glands in the skin secrete a slimy mucus over the skin, an adaptation that reduces drag during swimming. Like the ancient aquatic gnathostomes mentioned earlier, fishes have a lateral line system, which is evident as a row of tiny pits in the skin on either side of the body.

The details of fish reproduction vary extensively. Most species are oviparous, reproducing by external fertilization after the female sheds large numbers of small eggs. However, internal fertilization and birthing characterize other species.

Ray-Finned Fishes

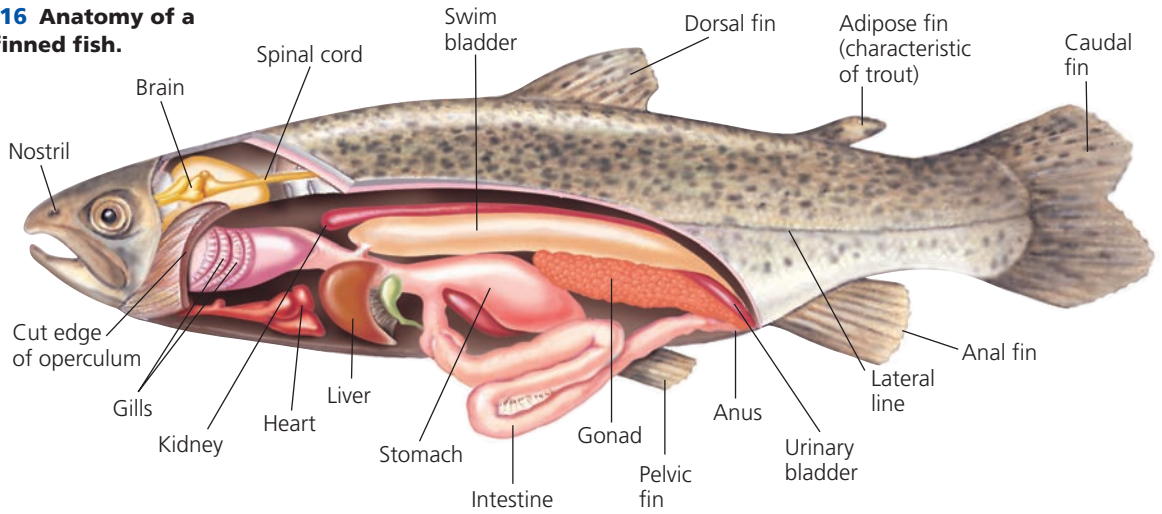
Nearly all the aquatic osteichthyans familiar to us are among the over 27,000 species of **ray-finned fishes** (Actinopterygii) **(Figure 34.17)**. Named for the bony rays that support their fins, the ray-finned fishes originated during the Silurian period (444–419 million years ago). The group has diversified greatly since that time, resulting in numerous species and many modifications in body form and fin structure that affect maneuvering, defense, and other functions.

Ray-finned fishes serve as a major source of protein for humans, who have harvested them for thousands of years. However, industrial-scale fishing operations appear to have driven some of the world's biggest fisheries to collapse. For example, after decades of abundant harvests, in the 1990s the catch of cod (*Gadus morhua*) in the northwest Atlantic plummeted to just 5% of its historic maximum, bringing cod fishing there to a near halt. Despite ongoing restrictions on the fishery, cod populations have yet to recover to sustainable levels. Ray-finned fishes also face other pressures from humans, such as the diversion of rivers by dams. Changing water flow patterns can hamper the fishes' ability to obtain food and interferes with migratory pathways and spawning grounds.

Lobe-Fins

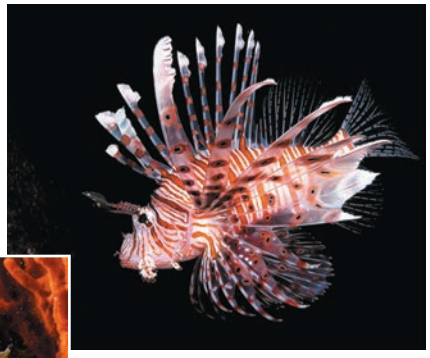
Like the ray-finned fishes, the other major lineage of osteichthyans, the **lobe-fins** (Sarcopterygii), also originated during

Figure 34.16 Anatomy of a trout, a ray-finned fish.



▲ Yellowfin tuna (*Thunnus albacares*) is a fast-swimming, schooling fish that is commercially important worldwide.

▶ Native to coral reefs of the Pacific Ocean, the brightly colored red lionfish (*Pterois volitans*) can inject venom through its spines, causing a severe and painful reaction in humans.



▲ The sea horse has a highly modified body form, as exemplified by *Hippocampus ramulosus,* shown above. Sea horses are unusual among animals in that the male carries the young during their embryonic development.

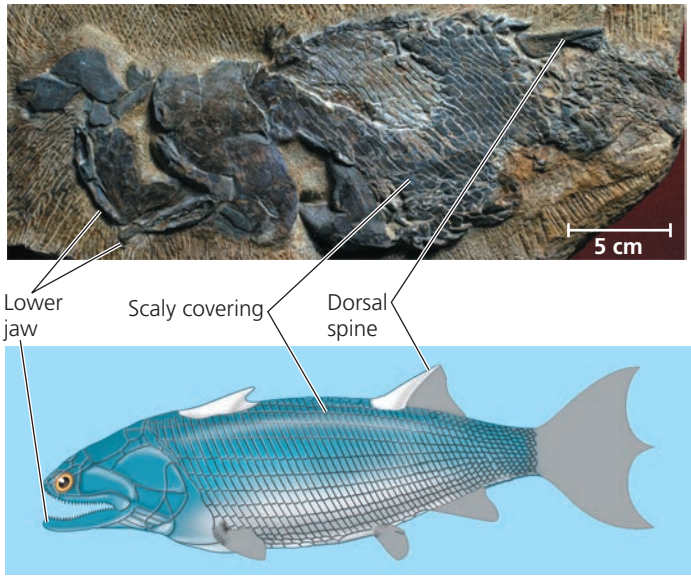
▲ The fine-spotted moray eel (*Gymnothorax dovii*) is a predator that ambushes prey from crevices in its coral reef habitat.

▲ Figure 34.17 Ray-finned fishes (Actinopterygii).

Video: Sea Horse Camouflage

the Silurian period **(Figure 34.18)**. The key derived character of lobe-fins is the presence of rod-shaped bones surrounded by a thick layer of muscle in their pectoral and pelvic fins. During the Devonian (419–359 million years ago), many lobe-fins lived in brackish waters, such as in coastal wetlands. There they may have used their lobed fins to help them move across logs or the muddy bottom (as do some living lobe-fins). Some Devonian lobe-fins were gigantic predators. It is not uncommon to find spike-shaped fossils of Devonian lobe-fin teeth as big as your thumb.

By the end of the Devonian period, lobe-fin diversity was dwindling, and today only three lineages survive. One lineage,

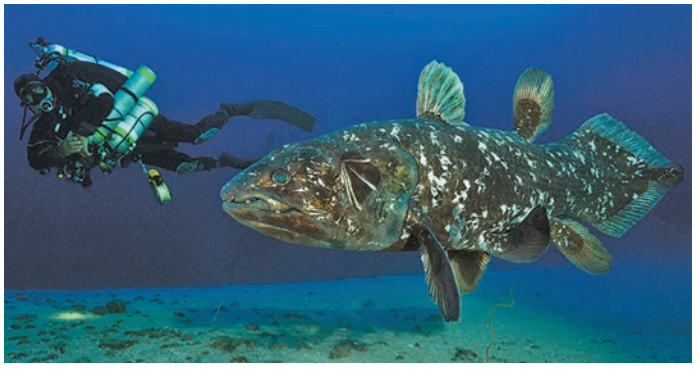


▲ Figure 34.18 A reconstruction of an ancient lobe-fin. Discovered in 2009, *Guiyu oneiros* is the earliest known lobe-fin, dating to 420 million years ago. The fossil of this species was nearly complete, allowing for an accurate reconstruction; regions shown in gray were missing from the fossil.

the coelacanths (Actinistia), was thought to have become extinct 75 million years ago. However, in 1938, fishermen caught a living coelacanth off the east coast of South Africa **(Figure 34.19)**. Until the 1990s, all subsequent discoveries were near the Comoros Islands in the western Indian Ocean. Since 1999, coelacanths have also been found at various places along the eastern coast of Africa and in the eastern Indian Ocean, near Indonesia. The Indonesian population may represent a second species.

The second lineage of living lobe-fins, the lungfishes (Dipnoi), is represented today by six species in three genera, all of which are found in the Southern Hemisphere. Lungfishes arose in the ocean but today are found only in fresh water, generally in stagnant ponds and swamps. They surface to gulp air into lungs connected to their pharynx. Lungfishes also have gills, which are the main organs for gas exchange in Australian lungfishes. When ponds shrink during the dry season, some lungfishes can burrow into the mud and estivate (wait in a state of torpor; see Concept 40.4).

▼ Figure 34.19 A coelacanth (*Latimeria*). These lobe-fins were found living off the coasts of southern Africa and Indonesia.



The third lineage of lobe-fins that survives today is far more diverse than the coelacanths or the lungfishes. During the mid-Devonian, these organisms adapted to life on land and gave rise to vertebrates with limbs and feet, called tetrapods—a lineage that includes humans.

CONCEPT CHECK 34.3

1. What derived characters do sharks and tuna share? What features distinguish tuna from sharks?
2. Describe key adaptations of aquatic gnathostomes.
3. **DRAW IT** ➤ Redraw Figure 34.2 to show four lineages: cyclostomes, lancelets, gnathostomes, and tunicates. Label the vertebrate common ancestor and circle the lineage that includes humans.
4. **WHAT IF?** ➤ Imagine that we could replay the history of life. Is it possible that a group of vertebrates that colonized land could have arisen from aquatic gnathostomes other than the lobe-fins? Explain.

For suggested answers, see Appendix A.

CONCEPT 34.4

Tetrapods are gnathostomes that have limbs

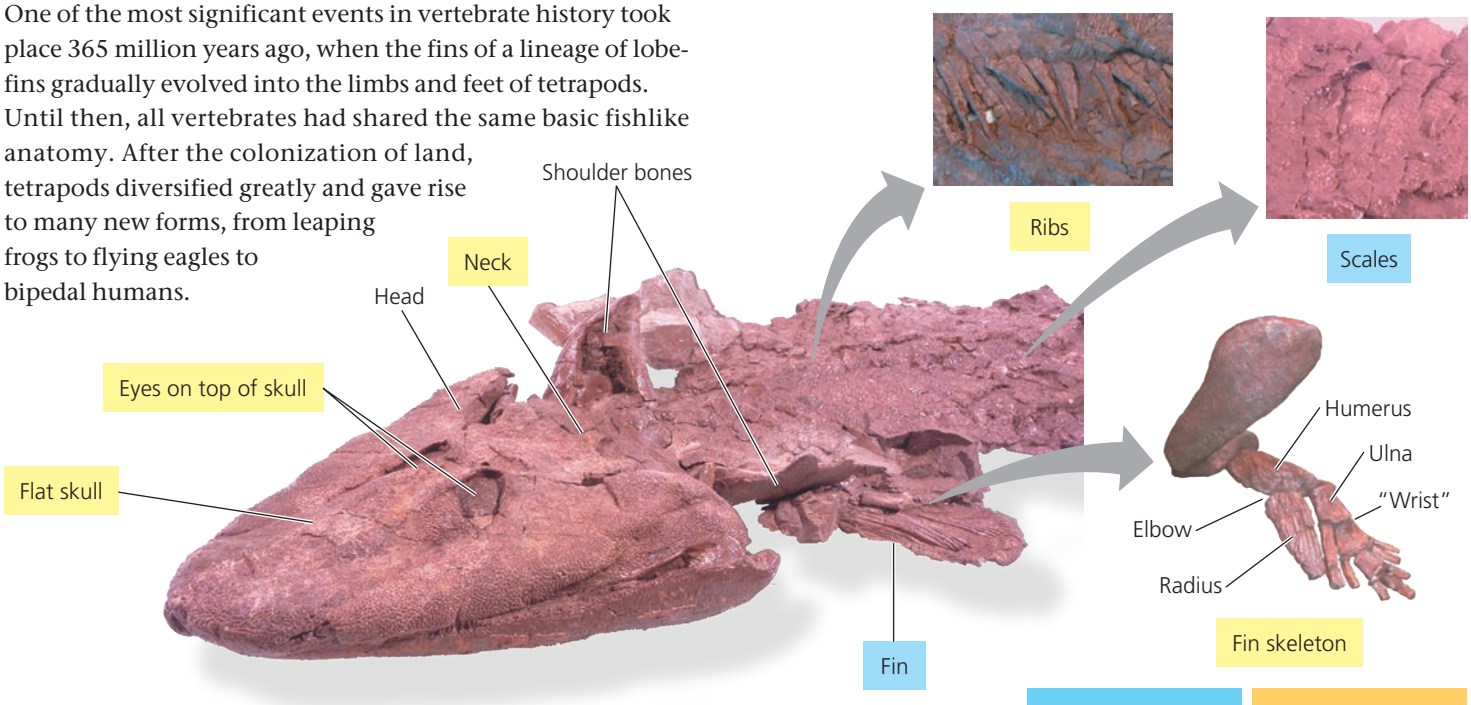
One of the most significant events in vertebrate history took place 365 million years ago, when the fins of a lineage of lobe-fins gradually evolved into the limbs and feet of tetrapods. Until then, all vertebrates had shared the same basic fishlike anatomy. After the colonization of land, tetrapods diversified greatly and gave rise to many new forms, from leaping frogs to flying eagles to bipedal humans.

Derived Characters of Tetrapods

The most significant character of **tetrapods** gives the group its name, which means "four feet" in Greek. In place of pectoral and pelvic fins, tetrapods have limbs with digits. Limbs support a tetrapod's weight on land, while feet with digits efficiently transmit muscle-generated forces to the ground when it walks.

Life on land selected for numerous other changes to the tetrapod body plan. In tetrapods, the head is separated from the body by a neck that originally had one vertebra on which the skull could move up and down. Later, with the origin of a second vertebra in the neck, the head could also swing from side to side. The bones of the pelvic girdle, to which the hind legs are attached, are fused to the backbone, permitting forces generated by the hind legs against the ground to be transferred to the rest of the body. Except for some fully aquatic species (such as the axolotl discussed below), the adults of living tetrapods do not have gills; during embryonic development, the pharyngeal clefts instead give rise to parts of the ears, certain glands, and other structures.

We'll discuss later how some of these characters were dramatically altered or lost in various lineages of tetrapods. In birds, for example, the pectoral limbs became wings, and in whales, the entire body converged toward a fishlike shape.



Fish Characters	Tetrapod Characters
Scales	Neck
Fins	Ribs
Gills and lungs	Fin skeleton
	Flat skull
	Eyes on top of skull

Figure 34.20 Discovery of a "fishapod": *Tiktaalik*. Paleontologists were on the hunt for fossils that could shed light on the evolutionary origin of tetrapods. Based on the ages of previously discovered fossils, researchers were looking for a dig site with rocks about 365–385 million years old. Ellesmere Island, in the Canadian Arctic, was one of the few such sites that was also likely to contain fossils, because it was once a river. The search at this site was rewarded by the discovery of fossils of a 375-million-year-old lobe-fin, named *Tiktaalik*. As shown in the chart and photographs, *Tiktaalik* exhibits both fish and tetrapod characters.

MAKE CONNECTIONS ➤ *Describe how* Tiktaalik's *features illustrate Darwin's concept of descent with modification (see Concept 21.2).*

HHMI Video: Great Transitions: The Origin of Tetrapods

The Origin of Tetrapods

As you have read, the Devonian coastal wetlands were home to a wide range of lobe-fins. Those that entered shallow, oxygen-poor water could have used their lungs to breathe air. Some species probably used their stout fins to swim and "walk" underwater across the bottom (moving their fins in an alternating gait, as do some living lobe-fins). This suggests that the tetrapod body plan did not evolve "out of nowhere" but was simply a modification of a preexisting body plan.

The recent discovery of a fossil called *Tiktaalik* provided new details on how this process of modification occurred **(Figure 34.20)**. Like a fish, this species had fins, gills, and lungs, and its body was covered in scales. But unlike a fish, *Tiktaalik* had a full set of ribs that would have helped it breathe air and support its body. Also unlike a fish, *Tiktaalik* had a neck and shoulders, allowing it to move its head about. In addition, the bones of *Tiktaalik*'s front fin have the same basic pattern found in all limbed animals: one bone (the humerus), followed by two bones (the radius and ulna), followed by a group of small bones that comprise the wrist. Finally, a 2014 study found that *Tiktaalik*'s pelvis and rear fin were larger and more robust than those of a fish; the pelvis is the bony structure to which hind limbs are attached in tetrapods. Although it is unlikely that *Tiktaalik* could walk on land, the skeletal structure of its fins and pelvis suggests that it could prop itself up and walk in water on its fins. Since *Tiktaalik* predates the oldest known tetrapod, its features suggest that key "tetrapod" traits, such as a wrist, ribs, and a neck, were in fact ancestral to the tetrapod lineage.

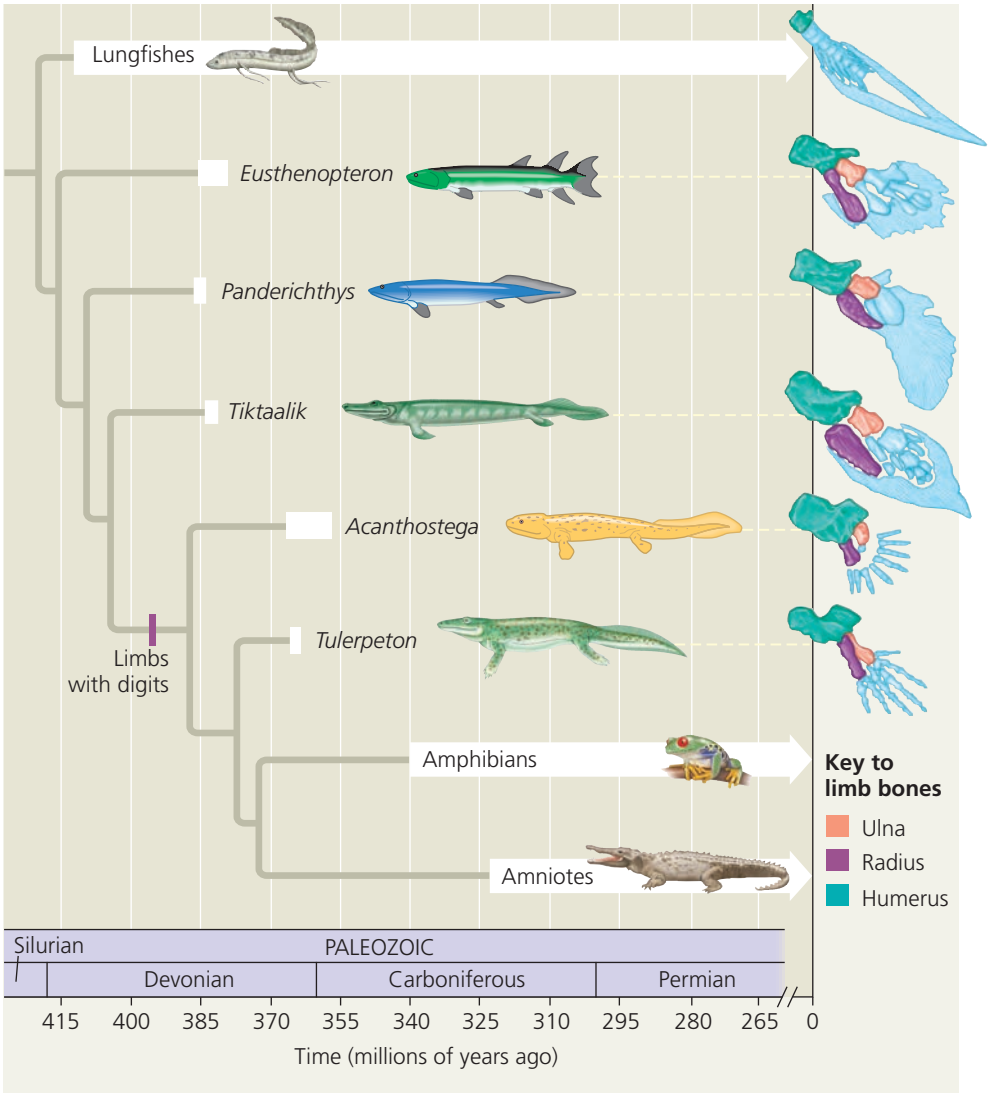
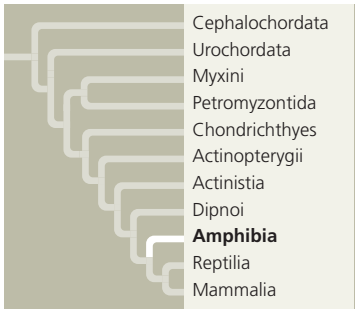


Figure 34.21 Steps in the origin of limbs with digits. The white bars on the branches of this diagram place known fossils in time; arrowheads indicate lineages that extend to today. The drawings of extinct organisms are based on fossilized skeletons, but the colors are fanciful.

WHAT IF? *If the most recent common ancestor of* Tulerpeton *and living tetrapods originated 370 million years ago, what range of dates would include the origin of amphibians?*

Tiktaalik and other extraordinary fossil discoveries have allowed paleontologists to reconstruct how fins became progressively more limb-like over time, culminating in the appearance in the fossil record of the first tetrapods 365 million years ago **(Figure 34.21)**. Over the next 60 million years, a great diversity of tetrapods arose. Some of these species retained functional gills and had weak limbs, while others had lost their gills and had stronger limbs that facilitated walking on land. Overall, judging from the morphology and locations of their fossils, most of these early tetrapods probably remained tied to water, a characteristic they share with some members of the most basal group of living tetrapods, the amphibians.

Amphibians



The **amphibians** are represented today by about 6,150 species in three clades: salamanders (clade Urodela, "tailed ones"), frogs (clade Anura, "tail-less ones"), and caecilians (clade Apoda, "legless ones").

Salamanders

There are about 550 known species of urodeles, or salamanders. Some are entirely aquatic, but others live on land as adults or throughout life. Most salamanders that live on land walk with a side-to-side bending of the body, a trait also found in early

(a) Order Urodela. Urodeles (salamanders) retain their tail as adults.

(b) Order Anura. Anurans (toads and frogs) lack a tail as adults.

(c) Order Apoda. Apodans, or caecilians, are legless, mainly burrowing amphibians.

Figure 34.22 Amphibians.

terrestrial tetrapods **(Figure 34.22a)**. Paedomorphosis is common among aquatic salamanders; the axolotl, for instance, retains larval features even when sexually mature (see Figure 25.24).

Frogs

Numbering about 5,420 species, anurans, or frogs, are better suited than salamanders to locomotion on land **(Figure 34.22b)**. Adult frogs use their powerful hind legs to hop along the terrain. Although often distinctive in appearance, the animals known as "toads" are simply frogs that have leathery skin or other adaptations for life on land. A frog nabs insects and other prey by flicking out its long, sticky tongue, which is attached to the front of the mouth. Frogs display a great variety of adaptations that help them avoid being eaten by larger predators. Their skin glands secrete distasteful or even poisonous mucus. Many poisonous species have color patterns that camouflage them or have bright coloration, which predators appear to associate with danger (see Figure 54.5).

Caecilians

The approximately 170 species of apodans, or caecilians, are legless and nearly blind, and superficially they resemble earthworms **(Figure 34.22c)**. Their absence of legs is a secondary adaptation, as they evolved from a legged ancestor. Caecilians inhabit tropical areas, where most species burrow in moist forest soil.

(a) The tadpole is an aquatic herbivore with a fishlike tail and internal gills.

(b) During metamorphosis, the gills and tail are resorbed, and walking legs develop. The adult frog will live on land.

(c) The adults return to water to mate. The male grasps the female, stimulating her to release eggs. The eggs are laid and fertilized in water. They have a jelly coat but lack a shell and would desiccate in air.

Figure 34.23 The "dual life" of a frog (*Rana temporaria*).

Lifestyle and Ecology of Amphibians

The term *amphibian* (derived from *amphibious*, meaning "both ways of life") refers to the life stages of many frog species that live first in water and then on land **(Figure 34.23)**. The larval stage of a frog, called a tadpole, is usually an aquatic herbivore with gills, a lateral line system resembling that of aquatic vertebrates, and a long, finned tail. The tadpole initially lacks legs; it swims by undulating its tail. During the metamorphosis that leads to the "second life," the tadpole develops legs, lungs, a pair of external eardrums, and a digestive system adapted to a carnivorous diet. At the same time, the gills disappear; the lateral line system also disappears in most species. The young frog crawls onto shore and becomes a terrestrial hunter. In spite of their name, however, many amphibians do not live a dual—aquatic and terrestrial—life. There are some strictly aquatic or strictly terrestrial frogs, salamanders, and caecilians. Moreover, salamander and caecilian larvae look much like the adults, and typically both the larvae and the adults are carnivorous.

Most amphibians are found in damp habitats such as swamps and rain forests. Even those adapted to drier habitats spend much of their time in burrows or under moist leaves, where humidity is high. One reason amphibians require relatively wet habitats is that they rely heavily on their moist skin for gas exchange—if their skin dries out, they cannot get enough oxygen. In addition, amphibians typically lay their eggs in water or in moist environments on land; their eggs lack a shell and dehydrate quickly in dry air.

Fertilization is external in most amphibians; the male grasps the female and spills his sperm over the eggs as the female sheds them (see Figure 34.23c). Some amphibian species lay vast numbers of eggs in temporary pools, and egg mortality is high. In contrast, other species lay relatively few eggs and display various types of parental care. Depending

Figure 34.24
A mobile nursery. A female marsupial frog (*Flectonotus fitzgeraldi*) incubates her eggs in pouches of skin on her back.

on the species, either males or females may house eggs on their back **(Figure 34.24)**, in their mouth, or even in their stomach. Certain tropical tree frogs stir their egg masses into moist, foamy nests that resist drying.

Many amphibians exhibit complex and diverse social behaviors, especially during breeding seasons. Frogs are usually quiet, but the males of many species vocalize to defend their breeding territory or to attract females. In some species, migrations to specific breeding sites may involve vocal communication, celestial navigation, or chemical signaling.

Over the past 30 years, zoologists have documented a rapid and alarming decline in amphibian populations in locations throughout the world. There appear to be several causes, including the spread of a disease-causing chytrid fungus (see Figure 31.25), habitat loss, climate change, and pollution. In some cases, declines have become extinctions. Recent studies indicate that at least 9 amphibian species have become extinct within the last four decades; more than 100 other species have not been observed in that time and are considered possibly extinct. In the **Problem-Solving Exercise**, you can explore one possible strategy to prevent amphibian deaths from fungal infections.

PROBLEM-SOLVING EXERCISE

Can declining amphibian populations be saved by a vaccine?

The Problem

Amphibian populations are declining rapidly worldwide. The fungus *Batrachochytrium dendrobatidis* (*Bd*) has contributed to this decline: This pathogen causes severe skin infections in many amphibian species, leading to massive die-offs. Efforts to save amphibians from *Bd* have met with limited success, and there is little evidence that frogs and other amphibians have acquired resistance to *Bd* on their own.

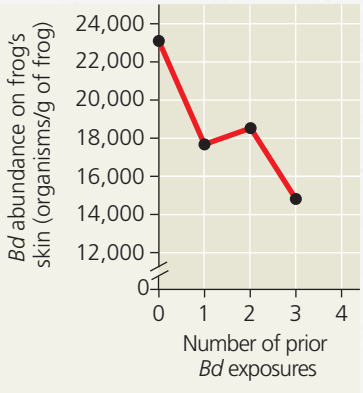
Yellow-legged frogs (*Rana muscosa*) in California killed by *Bd* infection

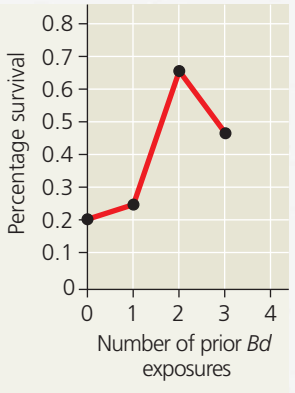
Instructors: A version of this Problem-Solving Exercise can be assigned in MasteringBiology.

In this exercise, you will investigate whether amphibians can acquire resistance to the fungal pathogen *Bd*.

Your Approach The principle guiding your investigation is that prior exposure to a pathogen can enable amphibians to acquire immunological resistance to that pathogen. To see whether this occurs after exposure to *Bd*, you will analyze data on acquired resistance in Cuban tree frogs (*Osteopilus septentrionalis*).

Your Data To create variation in number of prior exposures to *Bd*, Cuban tree frogs were exposed to *Bd* and cleared of their infection (using heat treatments) from 0 to 3 times; frogs with 0 prior exposures are referred to as "naive." Researchers then exposed frogs to *Bd* and measured mean abundance of *Bd* on the frog's skin, frog survival, and abundance of lymphocytes (a type of white blood cell involved in the immune response).





Number of prior *Bd* exposures	Thousands of lymphocytes per g frog
0	134
1	240
2	244
3	227

Your Analysis

1. Describe and interpret the results shown in the figure.
2. Graph the data in the table. Based on these data, develop a hypothesis that explains the results in the figure.
3. Breeding populations of amphibian species threatened by *Bd* have been established in captivity. In addition, evidence suggests that Cuban tree frogs can acquire resistance after exposure to dead *Bd*. Based on this information and your answers to questions 1 and 2, suggest a strategy for repopulating regions decimated by *Bd*.

CONCEPT CHECK 34.4

1. Describe the origin of tetrapods and identify some of their key derived traits.
2. Some amphibians never leave the water, whereas others can survive in relatively dry terrestrial environments. Contrast the adaptations that facilitate these two lifestyles.
3. **WHAT IF?** ➤ Scientists think that amphibian populations may provide an early warning system of environmental problems. What features of amphibians might make them particularly sensitive to environmental problems?

For suggested answers, see Appendix A.

CONCEPT 34.5

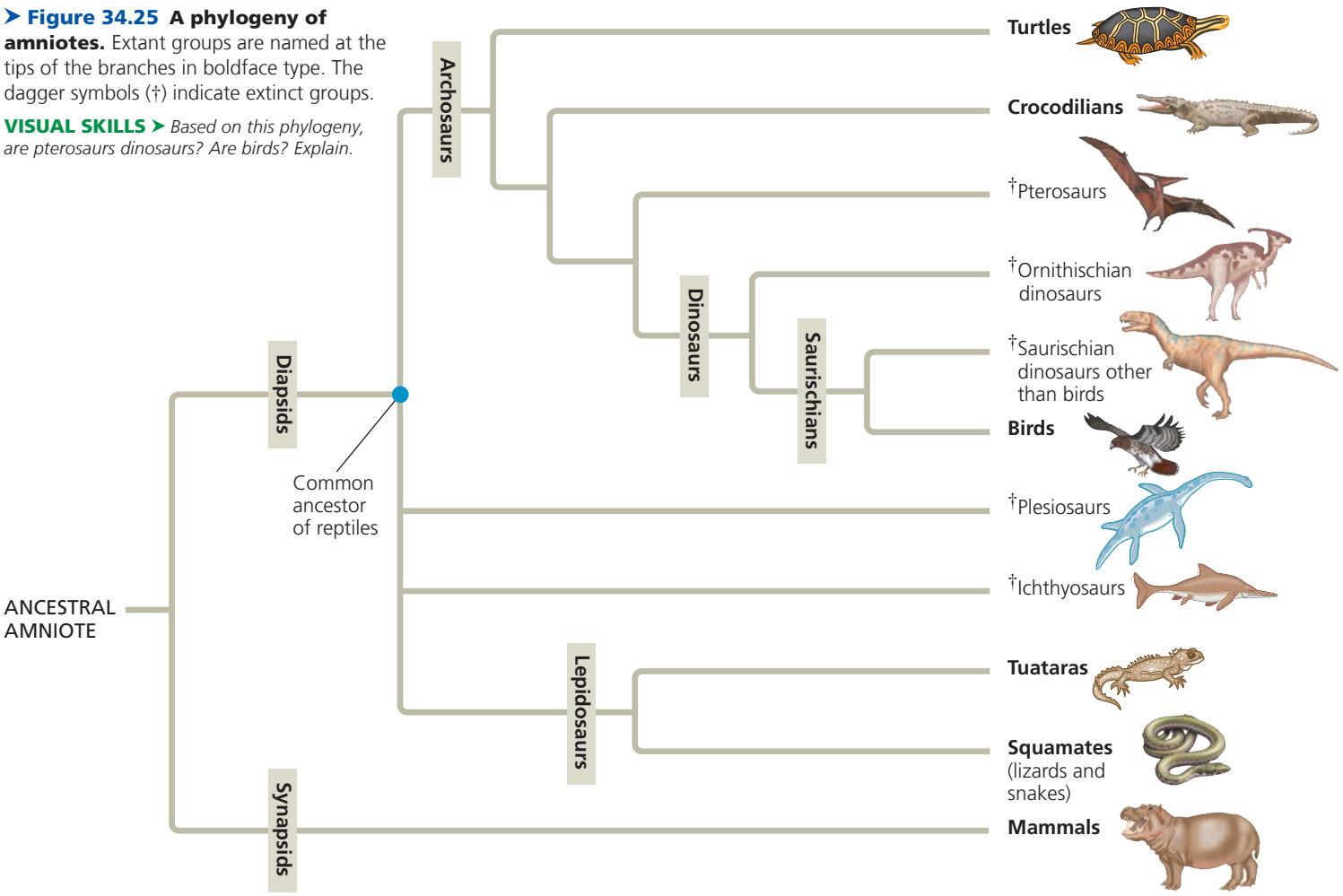
Amniotes are tetrapods that have a terrestrially adapted egg

The **amniotes** are a group of tetrapods whose extant members are the reptiles (including birds, as we'll discuss in this section) and mammals **(Figure 34.25)**. During their evolution, amniotes acquired a number of new adaptations to life on land.

Derived Characters of Amniotes

Amniotes are named for the major derived character of the clade, the **amniotic egg**, which contains four specialized membranes: the amnion, the chorion, the yolk sac, and the allantois **(Figure 34.26)**. Called *extraembryonic membranes* because they are not part of the body of the embryo itself, these membranes develop from tissue layers that grow out from the embryo. The amniotic egg is named for the amnion, which encloses a compartment of fluid that bathes the embryo and acts as a hydraulic shock absorber. The other membranes in the egg function in gas exchange, the transfer of stored nutrients to the embryo, and waste storage. The amniotic egg was a key evolutionary innovation for terrestrial life: It allowed the embryo to develop on land in its own private "pond," hence reducing the dependence of tetrapods on an aqueous environment for reproduction.

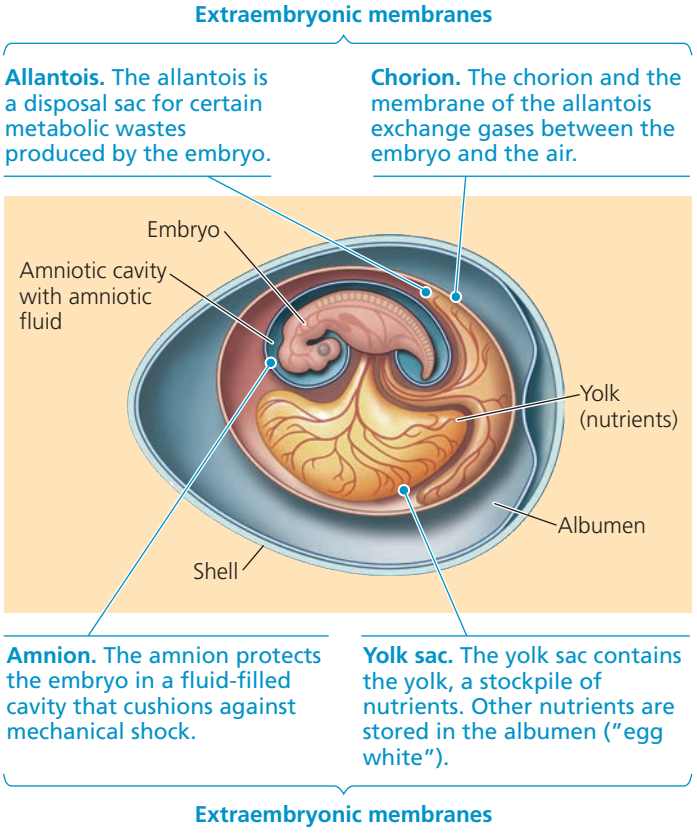
In contrast to the shell-less eggs of amphibians, the amniotic eggs of most reptiles and some mammals have a shell. A shell slows dehydration of the egg in air, an adaptation that helped amniotes to occupy a wider range of terrestrial habitats than amphibians, their closest living relatives. (Seeds played a similar role in the evolution of plants, as discussed in Concept 30.1.) Most mammals have lost the



➤ Figure 34.25 A phylogeny of amniotes. Extant groups are named at the tips of the branches in boldface type. The dagger symbols (†) indicate extinct groups.

VISUAL SKILLS ➤ *Based on this phylogeny, are pterosaurs dinosaurs? Are birds? Explain.*

Figure 34.26 The amniotic egg. The embryos of reptiles and mammals form four extraembryonic membranes: the allantois, chorion, amnion, and yolk sac. This diagram shows these membranes in the shelled egg of a reptile.



eggshell over the course of their evolution, and the embryo avoids desiccation by developing within the amnion inside the mother's body.

Amniotes have acquired other key adaptations to life on land. For example, amniotes use their rib cage to ventilate their lungs. This method is more efficient than throat-based ventilation, which amphibians use as a supplement to breathing through their skin. The increased efficiency of rib cage ventilation may have allowed amniotes to abandon breathing through their skin and develop less permeable skin, thereby conserving water.

Early Amniotes

The most recent common ancestor of living amphibians and amniotes lived 350 million years ago. No fossils of amniotic eggs have been found from that time, which is not surprising given how delicate they are. Thus, it is not yet possible to say when the amniotic egg evolved, although it must have existed in the last common ancestor of living amniotes, which all have amniotic eggs.

Based on where their fossils have been found, the earliest amniotes lived in warm, moist areas, as did the first tetrapods. Over time, early amniotes expanded into a wide range of new environments, including dry and high-latitude regions. Fossil evidence shows that the earliest amniotes resembled small

Figure 34.27 Artist's reconstruction of *Hylonomus*, an early amniote. About 25 cm long, this species lived 310 million years ago and probably ate insects and other small invertebrates.

lizards with sharp teeth, a sign that they were predators **(Figure 34.27)**. Later groups of amniotes also included herbivores, as evidenced by their grinding teeth and other features.

Reptiles

The **reptile** clade includes tuataras, lizards, snakes, turtles, crocodilians, and birds, along with a number of extinct groups, such as plesiosaurs and ichthyosaurs (see Figure 34.25).

As a group, the reptiles share several derived characters that distinguish them from other tetrapods. For example, unlike amphibians, reptiles have scales that contain the protein keratin (as does a human nail). Scales help protect the animal's skin from desiccation and abrasion. In addition, most reptiles lay their shelled eggs on land; the shell protects the egg from drying out **(Figure 34.28)**. Fertilization occurs internally, before the eggshell is secreted.

Reptiles such as lizards and snakes are sometimes described as "cold-blooded" because they do not use their metabolism extensively to control their body temperature. However, they do regulate their body temperature by using behavioral

Figure 34.28 Hatching reptiles. These baby panther chameleons (*Furcifer pardalis*) are breaking out of their parchment-like shells, a common type of shell among living reptiles other than birds.

adaptations. For example, many lizards bask in the sun when the air is cool and seek shade when the air is too warm. A more accurate description of these reptiles is to say that they are **ectothermic**, which means that they absorb external heat as their main source of body heat. By warming themselves directly with solar energy rather than through the metabolic breakdown of food, an ectothermic reptile can survive on less than 10% of the food energy required by a mammal of the same size. But the reptile clade is not entirely ectothermic; birds are **endothermic**, capable of maintaining body temperature through metabolic activity.

The Origin and Evolutionary Radiation of Reptiles

Fossil evidence indicates that the earliest reptiles lived about 310 million years ago and resembled lizards. Like all reptiles today, these early reptiles were **diapsids**. A key derived character of diapsids is a pair of holes on each side of the skull, behind the eye sockets; muscles pass through these holes and attach to the jaw, controlling jaw movement.

The diapsids are composed of two main lineages. One lineage gave rise to the **lepidosaurs**, which include tuataras, lizards, and snakes. This lineage also produced some marine reptiles, including the giant mososaurs. Some of these marine species rivaled today's whales in length; all of them are extinct. The other main diapsid lineage, the **archosaurs**, produced the turtles, crocodilians, pterosaurs, and dinosaurs. Our focus here will be on extinct lineages of archosaurs; we'll discuss living reptiles shortly.

Pterosaurs, which originated in the late Triassic, were the first tetrapods to exhibit flapping flight. The pterosaur wing was completely different from the wings of birds and bats. It consisted of a collagen-strengthened membrane that stretched between the trunk or hind leg and a very long digit on the foreleg. The smallest pterosaurs were no bigger than a sparrow, and the largest had a wingspan of nearly 11 m. They appear to have converged on many of the ecological roles later played by birds; some were insect-eaters, others grabbed fish out of the ocean, and still others filtered small animals through thousands of fine needlelike teeth. But by 66 million years ago, pterosaurs had become extinct.

On land, the **dinosaurs** diversified into a vast range of shapes and sizes, from bipeds the size of a pigeon to 45-m-long quadrupeds with necks long enough to let them browse the tops of trees. One lineage of dinosaurs, the ornithischians, were herbivores; they included many species with elaborate defenses against predators, such as tail clubs and horned crests. The other main lineage of dinosaurs, the saurischians, included the long-necked giants and a group called the **theropods**, which were bipedal carnivores. Theropods included the famous *Tyrannosaurus rex* as well as the ancestors of birds.

Dinosaurs once were considered slow, sluggish creatures. Since the 1970s, however, fossil discoveries and research have led to the conclusion that many dinosaurs were agile and fast moving. Dinosaurs had a limb structure that enabled them to walk and run more efficiently than could earlier tetrapods, which had a sprawling gait. Fossilized footprints and other evidence suggest that some species were social—they lived and traveled in groups, much as many mammals do today. Paleontologists have also discovered evidence that some dinosaurs built nests and brooded their eggs, as birds do today (see Figure 22.17). Finally, some anatomical evidence supports the hypothesis that at least some dinosaurs were endotherms.

All dinosaurs except birds became extinct by the end of the Cretaceous period (66 million years ago). Their extinction may have been caused at least in part by the asteroid or comet impact described in Concept 25.4. Some analyses of the fossil record are consistent with this idea in that they show a sudden decline in dinosaur diversity at the end of the Cretaceous. However, other analyses indicate that the number of dinosaur species had begun to decline several million years before the Cretaceous ended. Further fossil discoveries and new analyses will be needed to resolve this debate.

Next, we'll discuss the two extant lineages of reptiles, the lepidosaurs (tuataras, lizards, and snakes) and the archosaurs (turtles, crocodilians, and birds).

MB **Interview with Paul Sereno: Adventures hunting dinosaurs**
HHMI Video: How to Find a Dinosaur

Lepidosaurs

One surviving lineage of lepidosaurs is represented by two species of lizard-like reptiles called tuataras **(Figure 34.29a)**. Fossil evidence indicates that tuatara ancestors lived at least 220 million years ago. These organisms thrived on many continents well into the Cretaceous period and reached up to a meter in length. Today, however, tuataras are found only on 30 islands off the coast of New Zealand. When humans arrived in New Zealand 750 years ago, the rats that accompanied them devoured tuatara eggs, eventually eliminating the reptiles on the main islands. The tuataras that remain on the outlying islands are about 50 cm long and feed on insects, small lizards, and bird eggs and chicks. They can live to be

Figure 34.29 Extant reptiles (other than birds).

(a) Tuatara (*Sphenodon punctatus*)

over 100 years old. Their future survival depends on whether their remaining habitats are kept rat-free.

The other major living lineage of lepidosaurs consists of the lizards and snakes, or squamates, which number about 7,900 species **(Figure 34.29b** and **c)**. Many squamates are small; the Jaragua lizard, discovered recently in the Dominican Republic, is only 16 mm long—small enough to fit comfortably on a dime. In contrast, the Komodo dragon of Indonesia is a lizard that can reach a length of 3 m. It hunts deer and other large prey, delivering venom with its bite.

Snakes descended from lizards with legs—hence they are classified as legless lizards (see the opening paragraphs of Chapter 22). Today, some species of snakes retain vestigial pelvic and limb bones, providing evidence of their ancestry. Despite their lack of legs, snakes are quite proficient at moving on land, most often by producing waves of lateral bending that pass from head to tail. Force exerted by the bends against solid objects pushes the snake forward. Snakes can also move by gripping the ground with their belly scales at several points along the body while the scales at intervening points are lifted slightly off the ground and pulled forward.

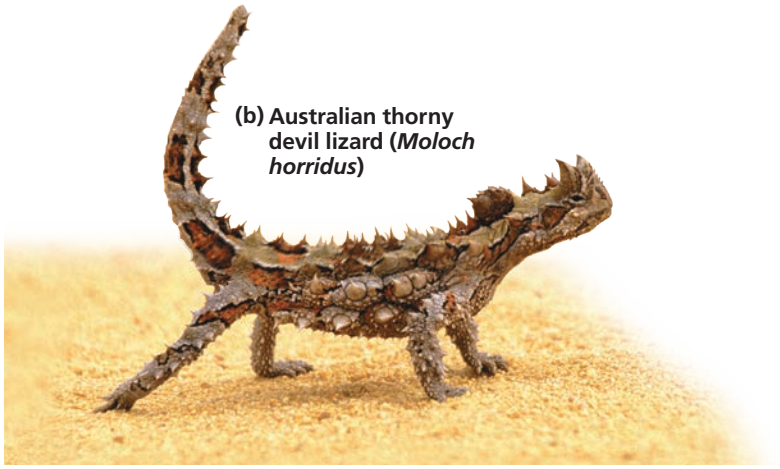
Snakes are carnivorous, and a number of adaptations aid them in hunting and eating prey. They have acute chemical sensors, and though they lack eardrums, they are sensitive to ground vibrations, which helps them detect the movements of prey. Heat-detecting organs between the eyes and nostrils of pit vipers, including rattlesnakes, are sensitive to minute temperature changes, enabling these night hunters to locate warm animals. Venomous snakes inject their toxin through a pair of sharp teeth that may be hollow or grooved. The flicking tongue is not venomous but helps fan odors toward olfactory (smell) organs on the roof of the mouth. Loosely articulated jawbones and elastic skin enable most snakes to swallow prey larger than the diameter of the snake's head (see Figure 23.14).

We'll conclude our survey of the reptiles by discussing the three clades of archosaurs with living members: The turtles, the crocodilians, and the birds.

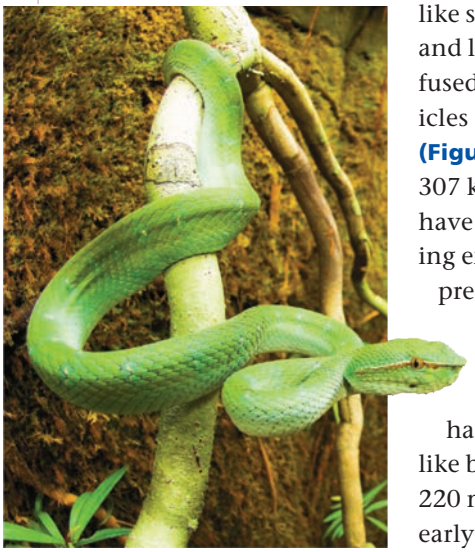
Turtles

Turtles are one of the most distinctive groups of reptiles alive today. For example, turtles do not have any holes in their skull behind the eye sockets, whereas other reptiles have two holes behind each eye socket. Recall that such skull holes are a key derived trait of the diapsids. Thus, until recently it was not clear whether turtles—like all other living reptiles—should be classified within the diapsid clade. However, in 2015, new fossil discoveries showed that early turtles had the skull openings found in other diapsids. This suggests that turtles are diapsids that have lost the holes in their skull over the course of evolution. The diapsid affinity of turtles was also confirmed by recent phylogenomic studies showing that turtles are archosaurs, more closely related to crocodilians and birds than to other reptiles (see Figure 34.25).

All turtles have a box-like shell made of upper and lower shields that are fused to the vertebrae, clavicles (collarbones), and ribs **(Figure 34.29d)**. Most of the 307 known species of turtles have a hard shell, providing excellent defense against predators. Fossil evidence shows that *Pappochelys*, a turtle that lived 240 million years ago, had a series of hard, shell-like bones along its belly. By 220 million years ago, another early turtle had a complete



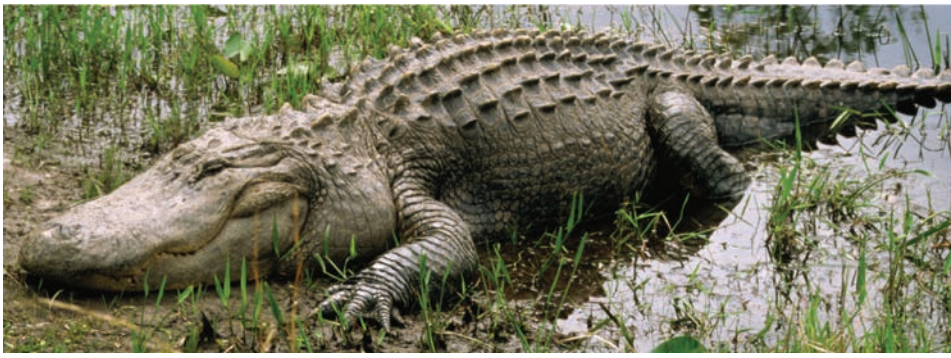
(b) Australian thorny devil lizard (*Moloch horridus*)



(c) Wagler's pit viper (*Tropidolaemus wagleri*)

(d) Black-breasted hill turtle (*Geomyda spengleri*)

(e) American alligator (*Alligator mississippiensis*)



lower shell but an incomplete upper shell, suggesting that turtles acquired full shells in stages.

The earliest turtles could not retract their head into their shell, but mechanisms for doing so evolved independently in two separate branches of turtles. The side-necked turtles fold their neck horizontally, while the vertical-necked turtles fold their neck vertically.

Some turtles have adapted to deserts, and others live almost entirely in ponds and rivers. Still others live in the sea. Sea turtles have a reduced shell and enlarged forelimbs that function as flippers. They include the largest living turtles, the deep-diving leatherbacks, which can exceed a mass of 1,500 kg and feed on jellies. Leatherbacks and other sea turtles are endangered by being caught in fishing nets, as well as by the residential and commercial development of the beaches where the turtles lay their eggs.

Crocodilians

Alligators and crocodiles (collectively called crocodilians) belong to a lineage that reaches back to the late Triassic. The earliest members of this lineage were small terrestrial quadrupeds with long, slender legs. Later species became larger and adapted to aquatic habitats, breathing air through their upturned nostrils. Some Mesozoic crocodilians grew as long as 12 m and may have attacked dinosaurs and other prey at the water's edge.

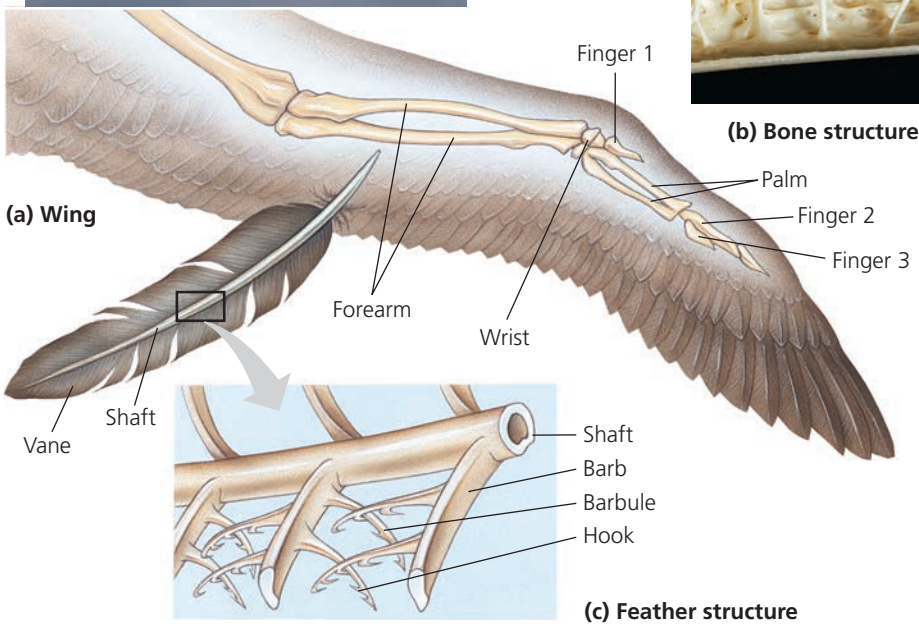
The 23 known species of living crocodilians are confined to warm regions of the globe. In the southeastern United States, the American alligator has made a comeback after spending years on the endangered species list.

Birds

There are about 10,000 species of birds in the world. Like crocodilians, birds are archosaurs, but almost every feature of their anatomy has been modified in their adaptation to flight.

Derived Characters of Birds Many of the characters of birds are adaptations that facilitate flight, including weight-saving modifications that make flying more efficient. For example, birds lack a urinary bladder, and the females of most species have only one ovary. The gonads of both females and males are usually small, except during the breeding season, when they increase in size. Living birds are also toothless, an adaptation that trims the weight of the head.

A bird's most obvious adaptations for flight are its wings and feathers **(Figure 34.30)**. Feathers are made of the protein β-keratin, which is also found in the scales of other reptiles. The shape and arrangement of the feathers form the wings into airfoils, and they illustrate some of the same principles of aerodynamics as the wings of an airplane. Power for flapping the wings comes from contractions of large pectoral (breast) muscles anchored to a keel on the sternum (breastbone). Some birds, such as eagles and hawks, have wings adapted for soaring on air currents and flap their wings only occasionally; other birds, including hummingbirds, must flap continuously to stay aloft (see Figure 34.34). Among the fastest birds are the appropriately named swifts, which can fly up to 170 km/hr.



(a) Wing

(c) Feather structure

(b) Bone structure

Figure 34.30 Form fits function: the avian wing and feather. **(a)** A wing is a remodeled version of the tetrapod forelimb. **(b)** The bones of many birds have a honeycombed internal structure and are filled with air. **(c)** A feather consists of a central air-filled shaft, from which radiate the vanes. The vanes are made up of barbs, which bear small branches called barbules. Birds have contour feathers and downy feathers. Contour feathers are stiff and contribute to the aerodynamic shapes of the wings and body. Their barbules have hooks that cling to barbules on neighboring barbs. When a bird preens, it runs the length of each contour feather through its beak, engaging the hooks and uniting the barbs into a precise shape. Downy feathers lack hooks, and the free-form arrangement of their barbs produces a fluffiness that provides insulation by trapping air.

MB Video: Soaring Hawk

Flight provides numerous benefits. It enhances scavenging and hunting, including enabling many birds to feed on flying insects, an abundant, nutritious food resource. Flight also provides ready escape from earthbound predators and enables some birds to migrate great distances to exploit different food resources and seasonal breeding areas.

Flying requires a great expenditure of energy from an active metabolism. Birds are endothermic; they use their own metabolic heat to maintain a high, constant body temperature. Feathers and in some species a layer of fat provide insulation that enables birds to retain body heat. The lungs have tiny tubes leading to and from elastic air sacs that improve airflow and oxygen uptake. This efficient respiratory system and a circulatory system with a four-chambered heart keep tissues well supplied with oxygen and nutrients, supporting a high rate of metabolism.

Flight also requires both acute vision and fine muscle control. Birds have color vision and excellent eyesight. The visual and motor areas of the brain are well developed, and the brain is proportionately larger than those of amphibians and nonbird reptiles.

Birds generally display very complex behaviors, particularly during breeding season, when they engage in elaborate courtship rituals. Because eggs have shells by the time they are laid, fertilization must be internal. Copulation usually involves contact between the openings to the birds' cloacas. After eggs are laid, the avian embryo must be kept warm through brooding by the mother, the father, or both, depending on the species.

The Origin of Birds Cladistic analyses of birds and reptilian fossils indicate that birds belong to the group of bipedal saurischian dinosaurs called theropods. Since the late 1990s, Chinese paleontologists have unearthed a spectacular trove of feathered theropod fossils that are shedding light on the origin of birds. Several species of dinosaurs closely related to birds had feathers with vanes, and a wider range of species had filamentous feathers. Such findings imply that feathers evolved long before powered flight. Among the possible functions of these early feathers were insulation, camouflage, and courtship display.

By about 160 million years ago, feathered theropods had evolved into birds. Many researchers consider *Archaeopteryx*, which was discovered in a German limestone quarry in 1861, to be the earliest known bird **(Figure 34.31)**. It had feathered wings but retained ancestral characters such as teeth, clawed digits in its wings, and a long tail. *Archaeopteryx* flew well at high speeds, but unlike a present-day bird, it could not take off from a standing position. Fossils of later birds from the Cretaceous show a gradual loss of certain ancestral dinosaur features, such as teeth and clawed forelimbs, as well as the acquisition of innovations found in extant birds, including a short tail covered by a fan of feathers.

Living Birds Clear evidence of Neornithes, the clade that includes the 28 orders of living birds, can be found before the Cretaceous-Paleogene boundary 66 million years ago. Several groups of living and extinct birds include one or more flightless species. The **ratites**, an order of birds that includes the ostrich, rhea, kiwi, cassowary, and emu, are all flightless **(Figure 34.32)**. In ratites, the sternal keel is absent, and the pectoral muscles are small relative to those of birds that can fly.

Penguins make up another flightless order of birds, but, like flying birds, they have powerful pectoral muscles. They use these muscles to "fly" in the water: As they swim, they flap their flipper-like wings in a manner that resembles the

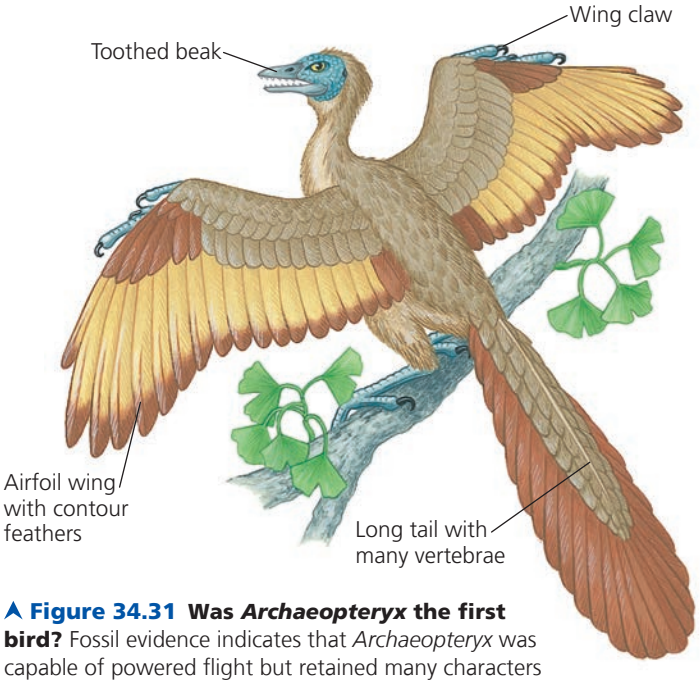


Figure 34.31 Was *Archaeopteryx* the first bird? Fossil evidence indicates that *Archaeopteryx* was capable of powered flight but retained many characters of nonbird dinosaurs. Although it has long been considered the first bird, recent fossil discoveries have sparked debate. Some analyses indicate that *Archaeopteryx* was a nonbird dinosaur closely related to the birds. Others indicate that *Archaeopteryx* was a bird—as traditionally thought—but that it was not the first bird.

HHMI Video: Great Transitions: The Origin of Birds

Figure 34.32 A kiwi (*Apteryx*), a flightless bird native to New Zealand.

Figure 34.33 A king penguin (*Aptenodytes patagonicus*) "flying" underwater. With their streamlined shape and powerful pectoral muscles, penguins are fast and agile swimmers.

flight stroke of a more typical bird **(Figure 34.33)**. Certain species of rails, ducks, and pigeons are also flightless.

Although the demands of flight have rendered the general body forms of many flying birds similar to one another, experienced bird-watchers can distinguish species by their profile, colors, flying style, behavior, and beak shape. The skeleton of a hummingbird's wing is unique, making it the only bird that can hover and fly backward **(Figure 34.34)**. Adult birds lack teeth, but during the course of avian evolution their beaks have taken on a variety of shapes suited to different diets. Some birds, such as parrots, have crushing beaks with which they can crack open hard nuts and seeds. Other birds, such as flamingoes, are filter feeders. Their beaks have "strainers" that enable them to capture food particles from the water **(Figure 34.35)**. Foot structure, too, shows considerable variation. Various birds use their feet for perching on branches **(Figure 34.36)**, grasping food, defense, swimming or walking, and even courtship (see Figure 24.3e).

Figure 34.34 Hummingbird feeding while hovering. A hummingbird can rotate its wings in all directions, enabling it to hover and fly backward.

Figure 34.35 A specialized beak. This greater flamingo (*Phoenicopterus ruber*) dips its beak into the water and strains out the food.

Figure 34.36 Feet adapted to perching. This great tit (*Parus major*) is a member of the Passeriformes, the perching birds. The toes of these birds can lock around a branch or wire, enabling the bird to rest for long periods.

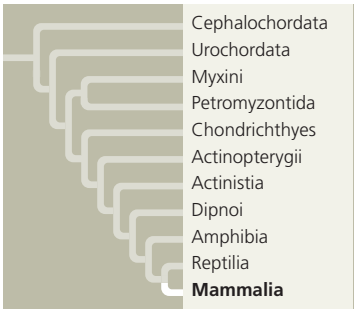
CONCEPT CHECK 34.5

1. Describe three key amniote adaptations for life on land.
2. Are snakes tetrapods? Explain.
3. Identify four avian adaptations for flight.
4. **VISUAL SKILLS ➤** Based on the phylogeny shown in Figure 34.25, identify the sister group for (a) reptiles; (b) squamates; and (c) the clade that includes crocodilians and birds.

For suggested answers, see Appendix A.

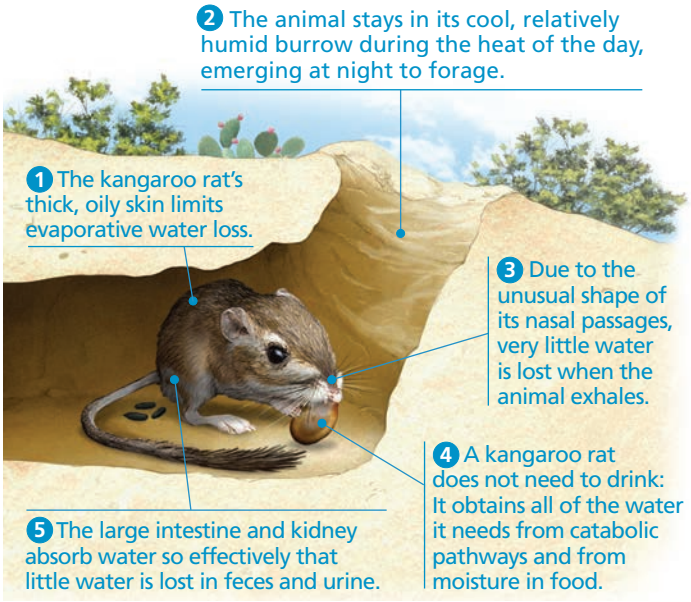
CONCEPT 34.6

Mammals are amniotes that have hair and produce milk



The reptiles we have been discussing represent one of the two living lineages of amniotes. The other amniote lineage is our own, the **mammals**. Today, there are more than 5,300 known species of mammals on Earth.

Figure 34.37 Adaptations of the kangaroo rat to its extremely dry habitat.



MAKE CONNECTIONS ➤ *Explain how the catabolic pathways mentioned in 4 could provide a kangaroo rat with water. (See Concept 10.1.)*

Derived Characters of Mammals

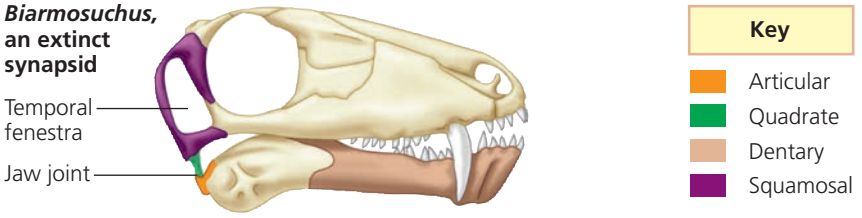
Mammals are named for their distinctive mammary glands, which produce milk for offspring. All mammalian mothers nourish their young with milk, a balanced diet rich in fats, sugars, proteins, minerals, and vitamins. Hair, another mammalian character, and a fat layer under the skin provide insulation that can conserve water and protect the body against extremes of heat or cold. Another mammalian adaptation for life on land is the kidney (see Figure 44.12), which is efficient at conserving water when removing wastes from the body. Some mammals, such as kangaroo rats, are so adept at conserving water that they can survive in arid environments while drinking little or no water at all **(Figure 34.37)**.

Like birds, mammals are endothermic, and most have a high metabolic rate. Efficient respiratory and circulatory systems (including a four-chambered heart) support a mammal's metabolism. Also as in birds, mammals generally have a larger brain than other vertebrates of equivalent size, and many species are capable learners. The relatively long duration of parental care extends the time for offspring to learn important survival skills by observing their parents. In addition, whereas the teeth of reptiles are generally uniform in size and shape, the jaws of mammals bear a variety of teeth with sizes and shapes adapted for chewing many kinds of foods. Humans, like most mammals, have teeth modified for shearing (incisors and canine teeth) and for crushing and grinding (premolars and molars).

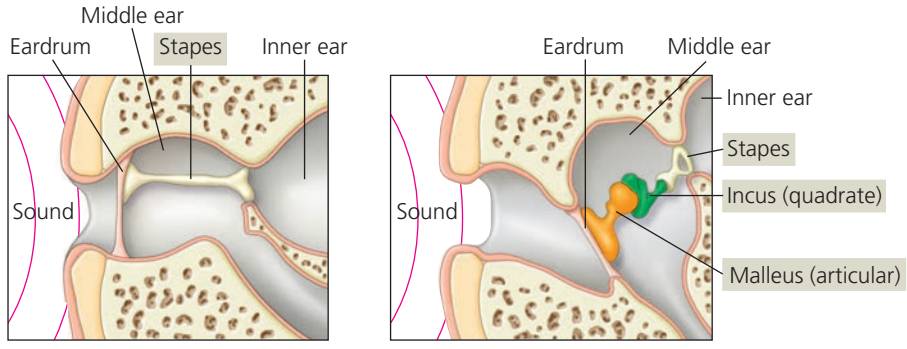
Early Evolution of Mammals

Mammals belong to a group of amniotes known as **synapsids**. Early nonmammalian synapsids lacked hair, had a sprawling gait, and laid eggs. A distinctive characteristic of synapsids is the single temporal fenestra, a hole behind the eye socket on each side of the skull. Humans retain this feature; your jaw muscles pass through the temporal fenestra and anchor on your temple. Fossil evidence shows that the jaw was remodeled as mammalian features arose gradually in successive lineages of earlier synapsids (see Figure 25.7); in all, these changes took more than 100 million years. In addition, two of the bones that formerly made up the jaw joint (the quadrate and the articular) were incorporated into the mammalian middle ear **(Figure 34.38)**. This evolutionary change is reflected in changes that occur during development. For example, as a mammalian embryo grows, the posterior region of its jaw—which in a reptile forms the articular bone—can be observed to detach from the jaw and migrate to the ear, where it forms the malleus.

Figure 34.38 The evolution of the mammalian ear bones. *Biarmosuchus* was a synapsid, a lineage that eventually gave rise to the mammals. Bones that transmit sound in the ear of mammals arose from the modification of bones in the jaw of nonmammalian synapsids.



(a) In *Biarmosuchus*, the meeting of the articular and quadrate bones formed the jaw joint.



(b) During the evolutionary remodeling of the mammalian skull, a new jaw joint formed between the dentary and squamosal bones (see Figure 25.7). No longer used in the jaw, the quadrate and articular bones became incorporated into the middle ear as two of the three bones that transmit sound from the eardrum to the inner ear.

MAKE CONNECTIONS ➤ *Review the definition of* exaptation *in Concept 25.6. Summarize the process by which exaptation occurs and explain how the incorporation of the articular and quadrate bones into the mammalian inner ear is an example.*

Synapsids evolved into large herbivores and carnivores during the Permian period (299–252 million years ago), and for a time they were the dominant tetrapods. However, the Permian-Triassic extinctions took a heavy toll on them, and their diversity fell during the Triassic (252–201 million years ago). Increasingly mammal-like synapsids emerged by the end of the Triassic. While not true mammals, these synapsids had acquired a number of the derived characters that distinguish mammals from other amniotes. They were small and probably hairy, and they likely fed on insects at night. Their bones show that they grew faster than other synapsids, suggesting that they probably had a relatively high metabolic rate; however, they still laid eggs.

During the Jurassic (201–145 million years ago), the first true mammals arose and diversified into many short-lived lineages. A diverse set of mammal species coexisted with dinosaurs in the Jurassic and Cretaceous periods, but these species were not abundant or dominant members of their communities, and most measured less than 1 m in length. One factor that may have contributed to their small size is that dinosaurs already occupied ecological niches of large-bodied animals.

By the early Cretaceous (140 million years ago), the three major lineages of mammals had emerged: those leading to monotremes (egg-laying mammals), marsupials (mammals with a pouch), and eutherians (placental mammals). After the extinction of large dinosaurs, pterosaurs, and marine reptiles during the late Cretaceous period, mammals underwent an adaptive radiation, giving rise to large predators and herbivores as well as flying and aquatic species.

Monotremes

Monotremes are found only in Australia and New Guinea and are represented by one species of platypus and four species of echidnas (spiny anteaters; **Figure 34.39**). Monotremes lay eggs, a character that is ancestral for amniotes and retained in most reptiles. Like all mammals, monotremes have hair and produce milk, but they lack nipples. Milk is secreted by glands on the belly of the mother. After hatching, the baby sucks the milk from the mother's fur.

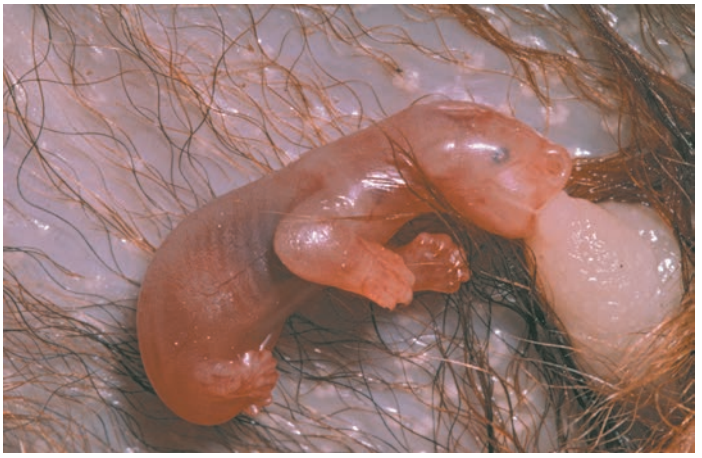
Figure 34.39 Short-beaked echidna (*Tachyglossus aculeatus*), an Australian monotreme. Monotremes have hair and produce milk, but they lack nipples. Monotremes are the only mammals that lay eggs (inset).

Marsupials

Opossums, kangaroos, and koalas are examples of the group called **marsupials**. Both marsupials and eutherians share derived characters not found among monotremes. They have higher metabolic rates and nipples that provide milk, and they give birth to live young. The embryo develops inside the uterus of the female's reproductive tract. The lining of the uterus and the extraembryonic membranes that arise from the embryo form a **placenta**, a structure in which nutrients diffuse into the embryo from the mother's blood.

A marsupial is born very early in its development and completes its embryonic development while nursing **(Figure 34.40a)**. In most species, the nursing young are held within a maternal pouch called a *marsupium*. A red kangaroo,

Figure 34.40 Australian marsupials.



(a) A young brushtail possum. The offspring of marsupials are born very early in their development. They finish their growth while nursing from a nipple (in their mother's pouch in most species).

(b) A greater bilby. The greater bilby is a digger and burrower that eats termites and other insects, along with the seeds, roots, and bulbs of various plants. The female's rear-opening pouch helps protect the young from dirt as the mother digs. Other marsupials, such as kangaroos, have a pouch that opens to the front.

for instance, is about the size of a honeybee at its birth, just 33 days after fertilization. Its back legs are merely buds, but its front legs are strong enough for it to crawl from the exit of its mother's reproductive tract to a pouch that opens to the front of her body, a journey that lasts a few minutes. In other species, the marsupium opens to the rear of the mother's body; in greater bilbies, this protects the young as their mother burrows in the dirt **(Figure 34.40b)**.

Marsupials existed worldwide during the Mesozoic era, but today they are found only in the Australian region and in North and South America. The biogeography of marsupials illustrates the interplay between biological and geologic evolution (see Concept 25.4). After the breakup of the supercontinent Pangaea, South America and Australia became island continents, and their marsupials diversified in isolation from the eutherians that began an adaptive radiation on the northern continents. Australia has not been in contact with another continent since early in the Cenozoic era, 66 million years ago. In Australia, convergent evolution has resulted in a diversity of marsupials that resemble eutherians in similar ecological roles in other parts of the world **(Figure 34.41)**. In contrast, although South America had a diverse marsupial fauna throughout the Paleogene, it has experienced several immigrations of eutherians. One of the most important occurred about 3 million years ago, when North and South America joined at the Panamanian isthmus and extensive two-way traffic of animals took place over the land bridge. Today, only three families of marsupials live outside the Australian region, and the only marsupials found in the wild in North America are a few species of opossum.

Eutherians (Placental Mammals)

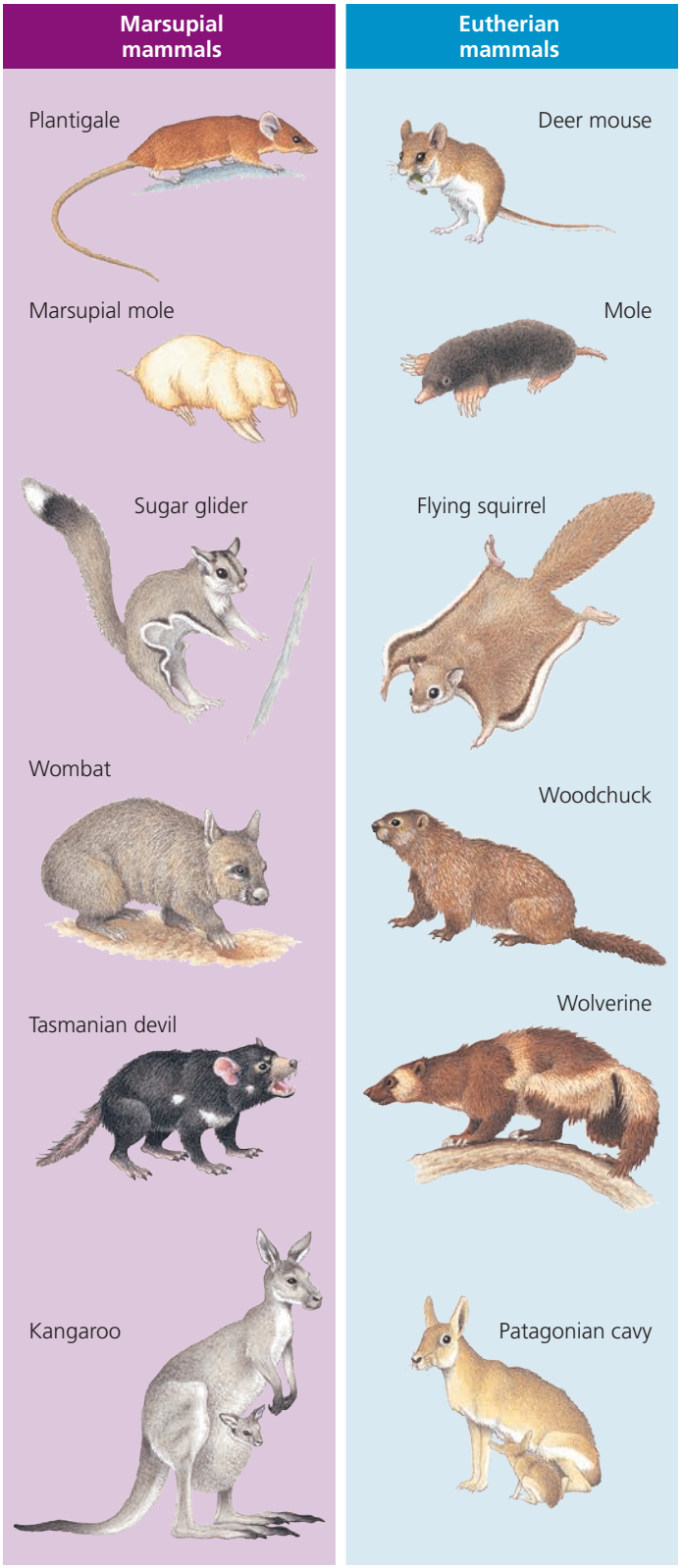
Eutherians are commonly called placental mammals because their placentas are more complex than those of marsupials. Eutherians have a longer pregnancy than marsupials. Young eutherians complete their embryonic development within the uterus, joined to their mother by the placenta. The eutherian placenta provides an intimate and long-lasting association between the mother and her developing young.

The major groups of living eutherians are thought to have diverged from one another in a burst of evolutionary change. The timing of this burst is uncertain: Molecular data suggest it occurred about 100 million years ago, while morphological data suggest it was about 60 million years ago. **Figure 34.42** explores several major eutherian orders and their phylogenetic relationships with each other as well as with the monotremes and marsupials.

Primates

The mammalian order Primates includes the lemurs, tarsiers, monkeys, and apes. Humans are members of the ape group.

Figure 34.41 Convergent evolution of marsupials and eutherians (placental mammals). (Drawings not to scale.)

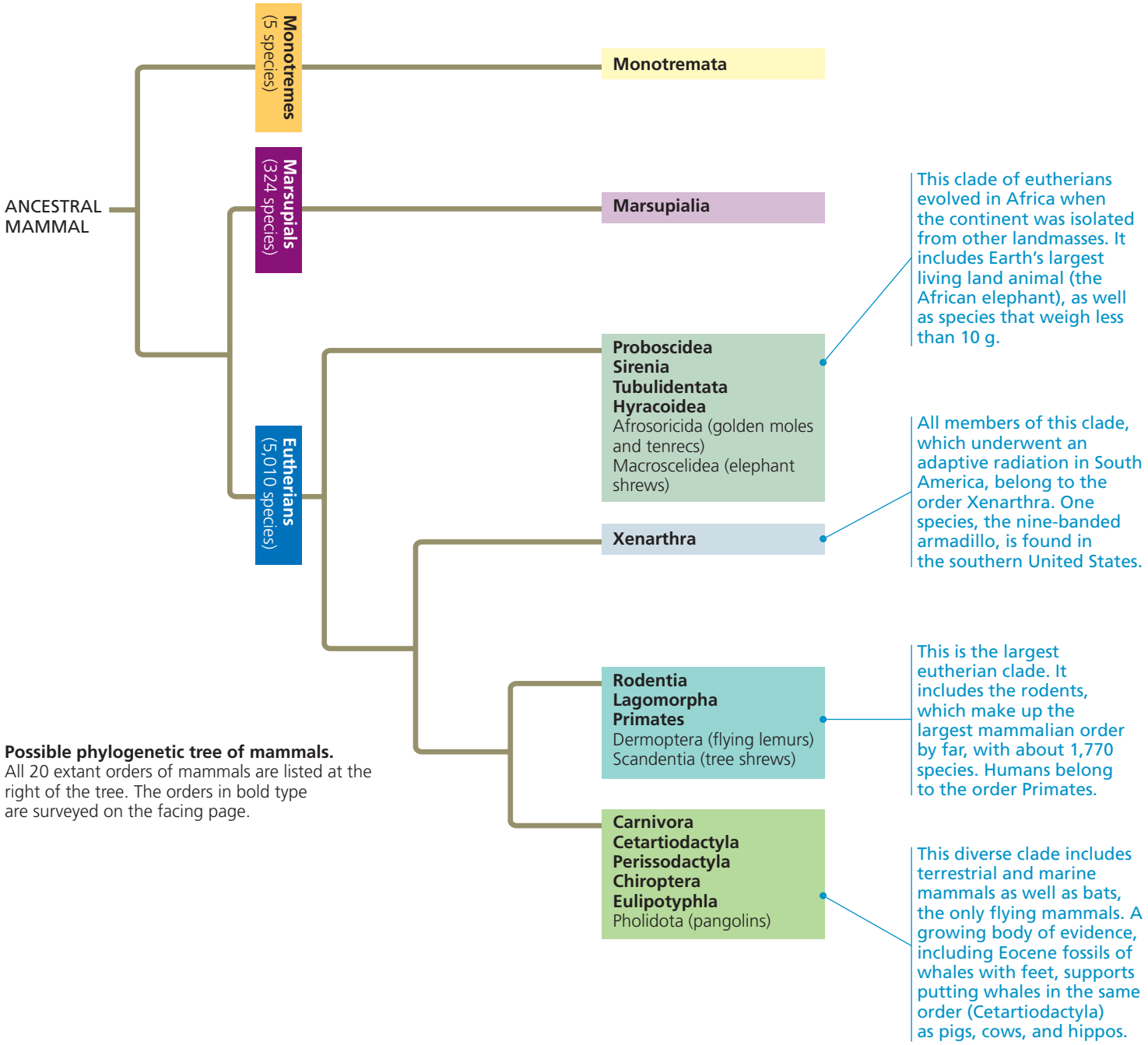


Derived Characters of Primates Most primates have hands and feet adapted for grasping, and their digits have flat nails instead of the narrow claws of other mammals. There are other characteristic features of the hands and feet, too, such as skin ridges on the fingers (which account for human

Figure 34.42 Exploring Mammalian Diversity

Phylogenetic Relationships of Mammals

Evidence from numerous fossils and molecular analyses indicates that monotremes diverged from other mammals about 180 million years ago and that marsupials diverged from eutherians (placental mammals) about 140 million years ago. Molecular systematics has helped to clarify the evolutionary relationships between the eutherian orders, though there is still no broad consensus on a phylogenetic tree. One current hypothesis, represented by the tree shown below, clusters the eutherian orders into four main clades.



Possible phylogenetic tree of mammals. All 20 extant orders of mammals are listed at the right of the tree. The orders in bold type are surveyed on the facing page.

Orders and Examples	Main Characteristics
Monotremata Platypuses, echidnas  Echidna	Lay eggs; no nipples; young suck milk from fur of mother
Proboscidea Elephants  African elephant	Long, muscular trunk; thick, loose skin; upper incisors elongated as tusks
Sirenia Manatees, dugongs  Manatee	Aquatic; finlike forelimbs and no hind limbs; herbivorous
Xenarthra Sloths, anteaters, armadillos  Tamandua	Reduced teeth or no teeth; herbivorous (sloths) or carnivorous (anteaters, armadillos)
Lagomorpha Rabbits, hares, picas  Jackrabbit	Chisel-like incisors; hind legs longer than forelegs and adapted for running and jumping; herbivorous
Carnivora Dogs, wolves, bears, cats, weasels, otters, seals, walruses  Coyote	Sharp, pointed canine teeth and molars for shearing; carnivorous
Cetartiodactyla Artiodactyls: sheep, pigs, cattle, deer, giraffes  Bighorn sheep	Hooves with an even number of toes on each foot; herbivorous
Cetaceans: whales, dolphins, porpoises  Pacific white-sided porpoise	Aquatic; streamlined body; paddle-like forelimbs and no hind limbs; thick layer of insulating blubber; carnivorous

Orders and Examples	Main Characteristics
Marsupialia Kangaroos, opossums, koalas  Koala	Completes embryonic development in pouch on mother's body
Tubulidentata Aardvarks  Aardvark	Teeth consisting of many thin tubes cemented together; eats ants and termites
Hyracoidea Hyraxes  Rock hyrax	Short legs; stumpy tail; herbivorous; complex, multi-chambered stomach
Rodentia Squirrels, beavers, rats, porcupines, mice  Red squirrel	Chisel-like, continuously growing incisors worn down by gnawing; herbivorous
Primates Lemurs, monkeys, chimpanzees, gorillas, humans  Golden lion tamarin	Opposable thumbs; forward-facing eyes; well-developed cerebral cortex; omnivorous
Perissodactyla Horses, zebras, tapirs, rhinoceroses  Indian rhinoceros	Hooves with an odd number of toes on each foot; herbivorous
Chiroptera Bats  Frog-eating bat	Adapted for flight; broad skinfold that extends from elongated fingers to body and legs; carnivorous or herbivorous
Eulipotyphla "Core insectivores": some moles, some shrews  Star-nosed mole	Eat mainly insects and other small invertebrates

fingerprints). Relative to other mammals, primates have a large brain and short jaws, giving them a flat face. Their forward-looking eyes are close together on the front of the face. Primates also exhibit relatively well-developed parental care and complex social behavior.

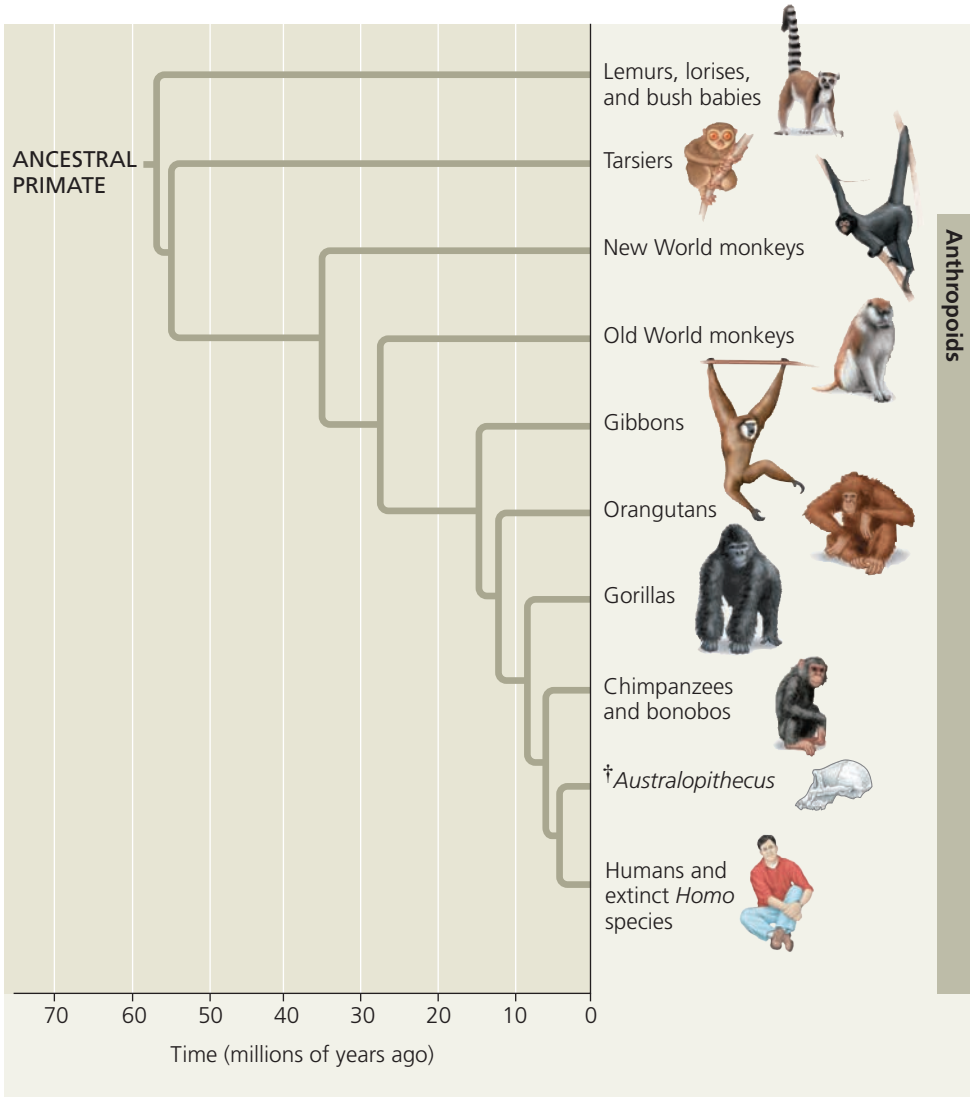
The earliest known primates were tree-dwellers, and many of the characteristics of primates are adaptations to the demands of living in the trees. Grasping hands and feet allow primates to hang onto tree branches. All living primates except humans have a big toe that is widely separated from the other toes, enabling them to grasp branches with their feet. All primates also have a thumb that is relatively movable and separate from the fingers, but monkeys and apes have a fully **opposable thumb**; that is, they can touch the ventral surface (fingerprint side) of the tip of all four fingers with the ventral surface of the thumb of the same hand. In monkeys and apes other than humans, the opposable thumb functions in a grasping "power grip." In humans, a distinctive bone structure at the base of the thumb allows it to be used for more precise manipulation. The unique dexterity of humans represents descent with modification from our tree-dwelling ancestors. Arboreal maneuvering also requires excellent eye-hand coordination. The overlapping visual fields of the two forward-facing eyes enhance depth perception, an obvious advantage when brachiating (traveling by swinging from branch to branch in trees).

➤ Figure 34.43 Verreaux's sifakas (*Propithecus verreauxi*), a type of lemur.

Living Primates

There are three main groups of living primates: (1) the lemurs of Madagascar **(Figure 34.43)** and the lorises and bush babies of tropical Africa and southern Asia; (2) the tarsiers, which live in southeastern Asia; and (3) the **anthropoids**, which include monkeys and apes and are found worldwide. The first group—lemurs, lorises, and bush babies—probably resemble early arboreal primates. The oldest known tarsier fossils date to 55 million years ago; along with DNA evidence, these fossils indicate that tarsiers are more closely related to anthropoids than to the lemur group **(Figure 34.44)**.

You can see in Figure 34.44 that monkeys do not form a clade but rather consist of two groups, the New and Old World monkeys. Both of these groups are thought to have originated in Africa or Asia. The fossil record indicates that New World monkeys first colonized South America roughly 25 million years ago. By that time, South America and Africa had

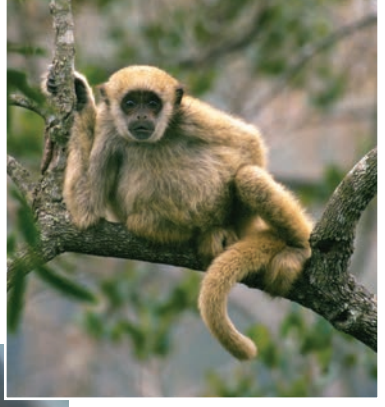


➤ Figure 34.44 A phylogenetic tree of primates. The fossil record indicates that the lineage leading to anthropoids diverged from other primates about 55 million years ago. New World monkeys, Old World monkeys, and apes (the clade that includes gibbons, orangutans, gorillas, chimpanzees, and humans) have been evolving as separate lineages for more than 25 million years. The lineage leading to humans and *Australopithecus* branched off from other apes between 6 and 7 million years ago.

VISUAL SKILLS ➤ *Is the phylogeny shown here consistent with the idea that humans evolved from chimpanzees? Explain.*

drifted apart, and monkeys may have reached South America from Africa by rafting on logs or other debris. What is certain is that New World monkeys and Old World monkeys underwent separate adaptive radiations during their many millions of years of separation **(Figure 34.45)**. All species of New World monkeys are arboreal, whereas Old World monkeys include ground-dwelling as well as arboreal species. Most monkeys in both groups are diurnal (active during the day) and usually live in bands held together by social behavior.

The other group of anthropoids consists of primates informally called apes **(Figure 34.46)**. The ape group includes the genera *Hylobates* (gibbons), *Pongo* (orangutans), *Gorilla* (gorillas), *Pan* (chimpanzees and bonobos), and *Homo* (humans). The apes diverged from Old World monkeys about 25–30 million years ago. Today, nonhuman apes are found exclusively in tropical regions of the Old World. With the exception of gibbons, living apes are larger than either New or Old World monkeys. All living apes have relatively long arms, short legs, and no tail. Although all nonhuman apes spend time in trees, only gibbons and orangutans are primarily arboreal. Social organization varies among the apes; gorillas and chimpanzees are highly social. Finally, compared to other



(a) New World monkeys, such as spider monkeys (shown here), squirrel monkeys, and capuchins, have a prehensile tail (one adapted for grasping) and nostrils that open to the sides.

(b) Old World monkeys lack a prehensile tail, and their nostrils open downward. This group includes macaques (shown here), mandrils, baboons, and rhesus monkeys.

Figure 34.45 New World monkeys and Old World monkeys.

Figure 34.46 Nonhuman apes.

(a) Gibbons, such as this Muller's gibbon, are found only in southeastern Asia. Their very long arms and fingers are adaptations for brachiating (swinging by the arms from branch to branch).

(b) Orangutans are shy apes that live in the rain forests of Sumatra and Borneo. They spend most of their time in trees; note the foot adapted for grasping and the opposable thumb.

(c) Gorillas are the largest apes; some males are almost 2 m tall and weigh about 200 kg. Found only in Africa, these herbivores usually live in groups of up to about 20 individuals.

(d) Chimpanzees live in tropical Africa. They feed and sleep in trees but also spend a great deal of time on the ground. Chimpanzees are intelligent, communicative, and social.

(e) Bonobos are in the same genus (*Pan*) as chimpanzees but are smaller. They survive today only in the African nation of Congo.

Video: Chimp Cracking Nut
Video: Gibbons Brachiating