science focus

Alternative mRNA Splicing in Disease

As you have read, the ability to combine the exons and introns of genes into new and novel combinations through alternative mRNA splicing is one of the mechanisms that allows humans to achieve a higher degree of complexity than simpler organisms without a huge increase in the number of genes. In more advanced organisms, the number of alternatively spliced mRNAs increases greatly. Recently, medical science has discovered that when this process goes awry, disease may result.

It has been known for many years that Gorlin syndrome, an autosomal dominant syndrome that includes aggressive skin cancer (Fig. 13A), multiple other tumors, either benign or cancerous, and cysts in various organs. Gorlin syndrome is linked to the tumor suppressor gene called *patched* located on chromosome 9. But a recent finding demonstrated that several new mutations in *patched* were not within the gene's exons, but within its introns! These mutations caused the mRNA to be spliced incorrectly, rendering the protein nonfunctional. Since at least 95% of most human genes consists of introns, there are likely many more such mutations to be discovered in other genetic disorders.

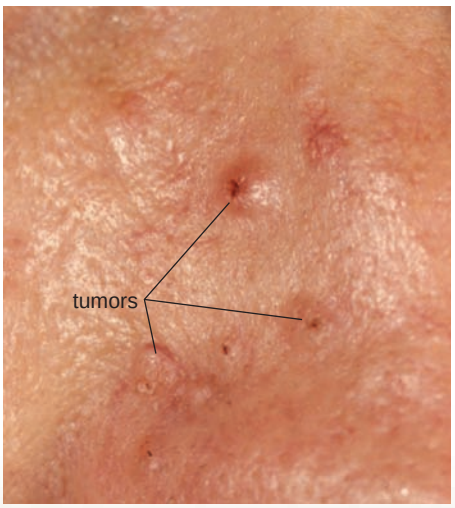


FIGURE 13A Skin tumors in patient with Gorlin syndrome.

Defective pre-mRNA splicing is also a major cause of spinal muscular atrophy (SMA), an autosomal recessive disorder that is a common cause of childhood mortality (Fig. 13B). Recent research shows that exon 7 of the *SMN2* gene tends to be left out of the mature mRNA in SMA patients, rendering the protein nonfunctional. The end result is a progressive loss of spinal cord motor neurons, and eventually paralysis and skeletal muscle atrophy. Scientists at Cold Spring Harbor Laboratories turned to antisense oligonucleotide technology, a relatively new technique, in an attempt to reverse the defect. The results were stunning—several of the oligonucleotides tested were able to promote the inclusion of exon 7 in the mature mRNA both in vitro and in cultured cells. These results raise the possibility that targeting aberrant pre-mRNA splicing may ultimately be a viable treatment for many disorders.

Alternative pre-mRNA splicing is also causing scientists to rethink strategies in disease treatment. For example, recent research indicates that the common drug acetaminophen actually targets an alternative version of the COX-1 protein in neurons. This protein variant arises from alternative splicing of the mRNA encoding COX-1 that only occurs in certain neurons. Ultimately, such new findings may allow investigators to design more powerful pain relievers with fewer and less severe side effects.

FIGURE 13B Spinal muscle atrophy.

Geneticists estimate that 80% or more of human genes undergo alternative mRNA splicing, and the estimate is constantly being revised upward. It is perhaps not surprising that this new frontier in gene regulation is redefining the standard approach to identifying the causes of illness and presenting new targets for the development of therapeutics.

Posttranslational Control

Posttranslational control begins once a protein has been synthesized and has become active. Posttranslational control represents the last chance a cell has for influencing gene expression.

Some proteins are not immediately active after synthesis. For example, at first bovine proinsulin is a single, long polypeptide that folds into a three-dimensional structure. Cleavage results in two smaller chains that are bonded together by disulfide (S—S) bonds. Only then is active insulin present. This ensures that some proteins only become active when it is appropriate for them to do so.

Some proteins are short-lived in cells because they are degraded or destroyed. For example, the cyclin proteins that control the cell cycle are only temporarily present. The cell has giant protein complexes called proteosomes that carry out the task of destroying proteins.

Check Your Progress 13.2

1. What are the five levels of genetic control in eukaryotes?
2. How does chromatin structure affect gene expression?
3. How does alternative processing of mRNA lead to genetic and phenotypic diversity?

13.3 Regulation Through Gene Mutations

A **gene mutation** is a permanent change in the sequence of bases in DNA. The effect of a DNA base sequence change on protein activity can range from no effect to complete inactivity. Germ-line mutations are those that occur in sex cells and can be passed to subsequent generations. Somatic mutations occur in body cells and, therefore, they may affect only a small number of cells in a tissue. Somatic mutations are not passed on to future generations, but they can lead to the development of cancer.

Causes of Mutations

Some mutations are spontaneous—they happen for no apparent reason—while others are induced by environmental influences. In most cases, **spontaneous mutations** arise as a result of abnormalities in normal biological processes. **Induced mutations** may result from exposure to toxic chemicals or radiation, which induce changes in the base sequence of DNA.

Spontaneous Mutations

Spontaneous mutations can be associated with any number of normal processes. For example, the movement of transposons from one chromosomal location to another can disrupt a gene and lead to an abnormal product. On rare occasions, a base in DNA can undergo a chemical change that leads to a miss pairing during replication. A subsequent base pair change may be carried forth in future generations. Spontaneous mutations due to DNA replication errors, however, are rare. DNA polymerase, the enzyme that carries out replication, proofreads the new strand against the old strand and detects any mismatched nucleotides, and each is usually replaced with a correct nucleotide. In the end, only about one mistake occurs for every 1 billion nucleotide pairs replicated.

Induced Mutations

Induced mutations are caused by **mutagens,** environmental factors that can alter the base composition of DNA. Among the best-known mutagens are radiation and organic chemicals. Many mutagens are also **carcinogens** (cancer-causing).

Chemical mutagens are present in many sources, including some of the food we eat and many industrial chemicals. The mutagenic potency of AF-2, a food additive once widely used in Japan, and of safrole, a flavoring agent once used to flavor root beer, caused them to be banned. Surprisingly, many naturally occurring substances like aflatoxin, produced in moldy grain and peanuts (and present in peanut butter at an average level of 2 parts per billion), and acrylamide, a natural product found in French fries, are also suspected mutagens.

Tobacco smoke contains a number of organic chemicals that are known carcinogens, and it is estimated that one-third of all cancer deaths can be attributed to smoking. Lung cancer is the most frequent lethal cancer in the United States, and smoking is also implicated in the development of cancers of the mouth, larynx, bladder, kidney, and pancreas. The greater the number of cigarettes smoked per day, the earlier the habit starts, and the higher the tar content, the greater the possibility of these cancers. When smoking is combined with drinking alcohol, the risk of these cancers increases even more.

Scientists use the Ames test for mutagenicity to hypothesize that a chemical can be carcinogenic (Fig. 13.10). In the Ames test, a histidine-requiring strain of bacteria is exposed to a chemical. If the chemical is mutagenic, the bacteria can grow without histidine. A large number of chemicals used in agriculture and industry give a positive Ames test. Examples are ethylene dibromide (EDB), which is added to leaded gasoline (to vaporize lead deposits in the engine and send them out the exhaust), and ziram, which is used to prevent fungus disease on crops. Some drugs, such as isoniazed (used to prevent tuberculosis), are mutagenic according to the Ames test.

Aside from chemicals, certain forms of radiation, such as X-rays and gamma rays, are called ionizing radiation

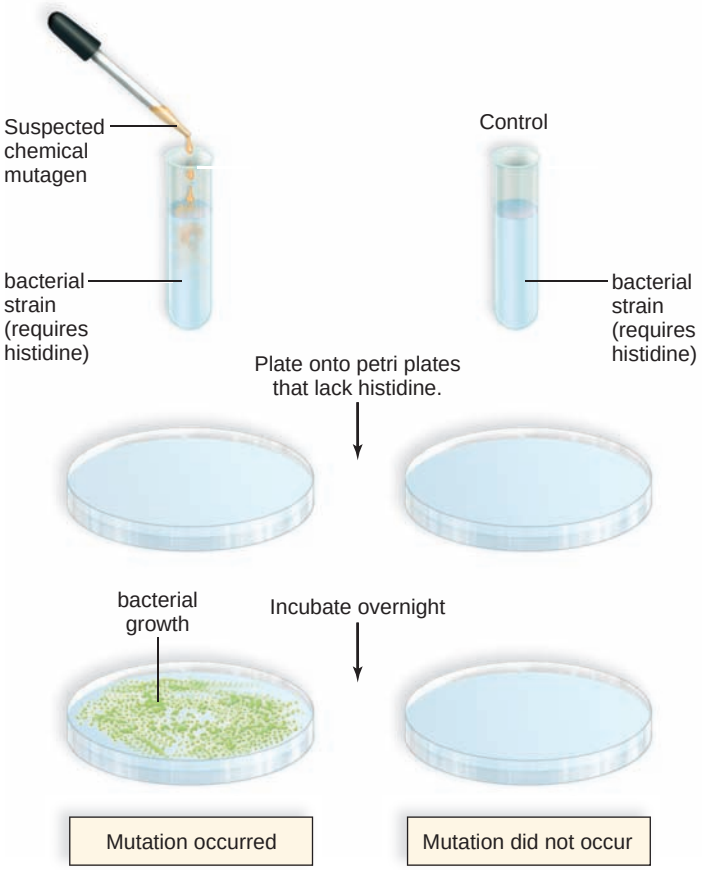


FIGURE 13.10 The Ames test for mutagenicity.

A bacterial strain that requires histidine as a nutrient is exposed to a suspected chemical mutagen, but a control is not exposed. The bacteria are plated on a medium that lacks histidine and only the bacteria exposed to the chemical show growth. A mutation allowed the bacteria to grow; therefore, the chemical can be carcinogenic.

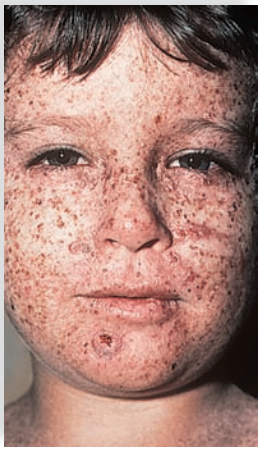


FIGURE 13.11
Xeroderma pigmentosum.

In xeroderma pigmentosum, deficient DNA repair enzymes leave the skin cells vulnerable to the mutagenic effects of ultraviolet light, allowing many induced mutations to accumulate. Hundreds of skin cancers (small dark spots) appear on the skin exposed to the sun. This individual also has a tumor on the bridge of the nose.

because they create free radicals, ionized atoms with unpaired electrons. Free radicals react with and alter the structure of other molecules, including DNA. Ultraviolet (UV) radiation is easily absorbed by the pyrimidines in DNA. Wherever there are two thymine molecules next to one another, ultraviolet radiation may cause them to bond together, forming *thymine dimers*. A kink results in the DNA. Usually, these dimers are removed by **DNA repair enzymes,** which constantly monitor DNA and fix any irregularities. One enzyme excises a portion of DNA that contains the dimer; another makes a new section by using the other strand as a template; and still another seals the new section in place. The importance of these repair enzymes is exemplified by individuals with the condition known as xeroderma pigmentosum. They lack some of the repair enzymes, and as a consequence, these individuals have a high incidence of skin cancer because of the large number of mutations that accumulate over time (Fig. 13.11). Also, repair enzymes can fail as when skin cancer develops because of excessive sunbathing or prolonged exposure to X-rays.

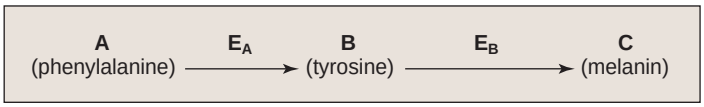
Effect of Mutations on Protein Activity

Point mutations involve a change in a single DNA nucleotide and, therefore, a possible change in a specific amino acid. The base change in the second row of Figure 13.12*a* has no effect on the resulting amino acid in hemoglobin; the change in the third row, however, codes for the amino acid glutamic acid instead of valine. This base change accounts for the genetic disorder sickle-cell disease because the incorporation of valine, instead of glutamic acid, causes hemoglobin molecules to form semirigid rods, and the red blood cells become sickle shaped. (Compare Figure 13.12*b* to Figure 13.12*c*.) Sickle-shaped cells clog blood vessels and die off more quickly than normal-shaped cells. The base change in the fourth row of Figure 13.12*a* may also have drastic results because the DNA now codes for a stop codon.

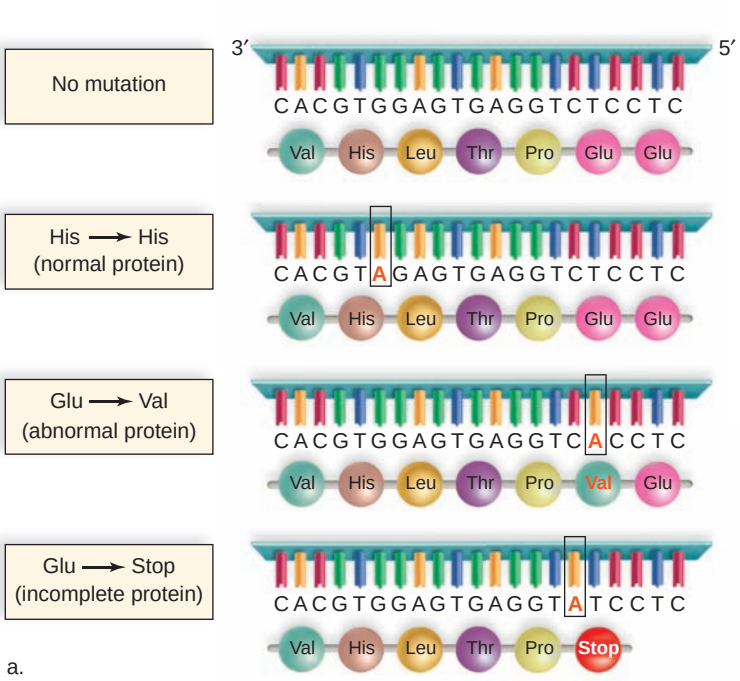
Frameshift mutations occur most often because one or more nucleotides are either inserted or deleted from DNA. The result of a frameshift mutation can be a completely new sequence of codons and nonfunctional protein. Here is how this occurs: The sequence of codons is read from a specific starting point, as in this sentence, THE CAT ATE THE RAT. If the letter C is deleted from this sentence and the reading frame is shifted, we read THE ATA TET HER AT—something that doesn't make sense.

Nonfunctional Proteins

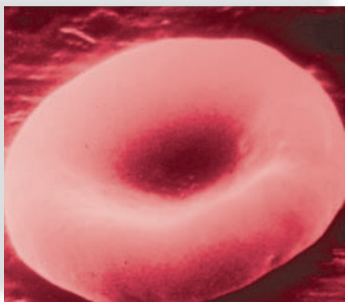
A single nonfunctioning protein can have a dramatic effect on the phenotype, because enzymes are often a part of metabolic pathways. One particular metabolic pathway in cells is as follows:



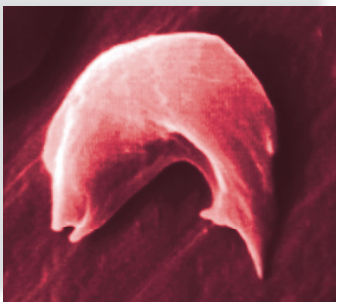
If a faulty code for enzyme E_A is inherited, a person is unable to convert the molecule A to B. Phenylalanine builds



a.



b. Normal red blood cell



c. Sickled red blood cell

FIGURE 13.12 **Point mutations in hemoglobin.**

The effect of a point mutation can vary. **a.** Starting at the *top:* Normal sequence of bases in hemoglobin; next, the base change has no effect; next, due to base change, DNA now codes for valine instead of glutamic acid, and the result is that normal red blood cells (**b**) become sickle shaped **c**; next, base change will cause DNA to code for termination and the protein will be incomplete.

up in the system, and the excess causes mental retardation and the other symptoms of the genetic disorder phenylketonuria (PKU). In the same pathway, if a person inherits a faulty code for enzyme E_B, then B cannot be converted to C, and the individual is an albino.

A rare condition called androgen insensitivity is due to a faulty receptor for androgens, which are male sex hormones such as testosterone. Although there is plenty of testosterone in the blood, the cells are unable to respond to it. Female instead of male external genitals form, and female instead of male secondary sex characteristics occur. The individual, who appears to be a normal female, may be prompted to seek medical advice when menstruation never occurs. The karyotype is that of a male rather than a female, and the individual does not have the internal sexual organs of a female.

Mutations Can Cause Cancer

It is estimated that one-third of the children born in 1999 will develop cancer at some time in their lives. Of these affected individuals, one-third of the females and one-fourth of the males will die due to cancer. In the United States, the three deadliest forms of cancer are lung cancer, colon and rectal cancer, and breast cancer.

The development of cancer involves a series of accumulating mutations that can be different for each type of cancer. As discussed in Chapter 9, tumor suppressor genes ordinarily act as brakes on cell division, especially when it begins to occur abnormally. Proto-oncogenes stimulate cell division but are usually turned off in fully differentiated nondividing cells. When proto-oncogenes mutate, they become oncogenes that are active all the time. Carcinogenesis begins with the loss of tumor suppressor gene activity and/or the gain of oncogene activity. When tumor suppressor genes are inactive and oncogenes are active, cell division occurs uncontrollably because a cell signaling pathway that reaches from the plasma membrane to the nucleus no longer functions as it should (Fig. 13.13 and Fig. 13.14).

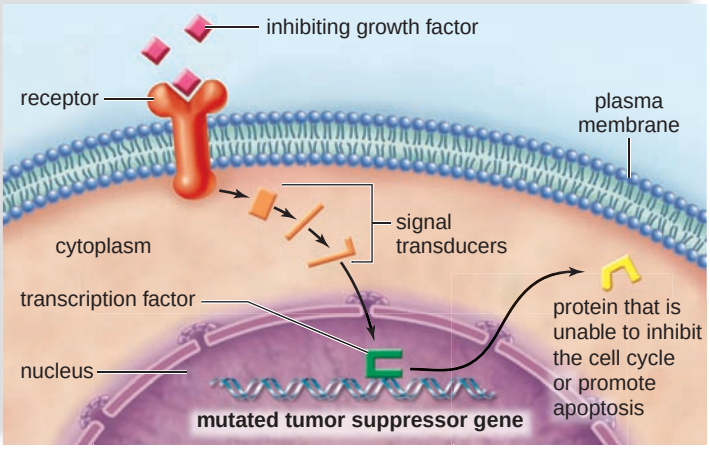


FIGURE 13.13 Cell signaling pathway that stimulates a mutated tumor suppressor gene.

A mutated tumor suppressor gene codes for a product that directly or indirectly stimulates the cell cycle.

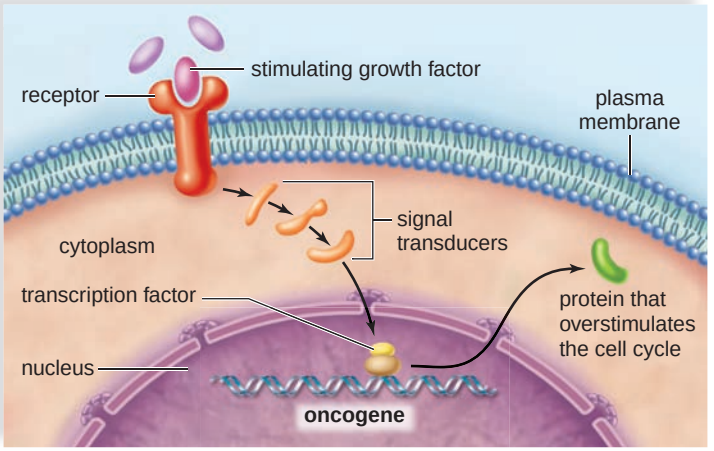


FIGURE 13.14 Cell signaling pathway that stimulates an oncogene.

An oncogene codes for a product that either directly or indirectly overstimulates the cell cycle.

It often happens that tumor suppressor genes and proto-oncogenes code for transcription factors or proteins that control transcription factors. As we have seen, transcription factors are a part of the rich and diverse types of mechanisms that control gene expression in cells. They are of fundamental importance to DNA replication and repair, cell growth and division, control of apoptosis, and cellular differentiation. Therefore, it is not surprising that inherited or acquired defects in transcription factor structure and function contribute to the development of cancer.

To take an example, a major tumor suppressor gene called *p53* is more frequently mutated in human cancers than any other known gene. It has been found that the p53 protein acts as a transcription factor, and as such is involved in turning on the expression of genes whose products are cell cycle inhibitors (see page 153). *p53* also promotes apoptosis (programmed cell death) when it is needed. The retinoblastoma protein (RB) controls the activity of a transcription factor for cyclin D and other genes whose products promote entry into the S stage of the cell cycle. When the tumor suppressor gene *p16* mutates, the RB protein is always available, and the result is too much active cyclin D in the cell.

Mutations in many other genes also contribute to the development of cancer. Several proto-oncogenes code for ras proteins, which are needed for cells to grow, to make new DNA, and to not grow out of control. A point mutation is sufficient to turn a normally functioning *ras* proto-oncogene into an oncogene. Abnormal growth results.

Check Your Progress 13.3

1. What are some common causes of spontaneous and induced mutations?
2. Explain how a frameshift mutation may disrupt a gene's function.

Connecting the Concepts

The characteristics of specialized cells, such as nerve cells, muscle fibers, and reproductive cells, are maintained by the differences in genes that each cell type expresses. Gene regulation determines whether a gene is expressed and/or the degree to which a gene is expressed. The collective work of many researchers was required to determine that the regulation of genes can occur at most stages during the processes of transcription and translation and that even the chromatin structure of the chromosome containing the gene can influence whether or not a gene is active.

Further control of gene expression, and thus the type of proteins produced by a cell, is achieved through mechanisms such as alternative pre-mRNA splicing, by influencing the rate at which an mRNA leaves the nucleus, and through regulation by miRNAs. Together, these mechanisms alter gene expression to suit an organism's needs.

All living things are subject to genetic mutations or changes in the base sequence of DNA. The effects of mutations vary greatly, but in some cases, mutations may cause genes to be abnormally turned off or on, or expressed in abnormal quantity. Such mutations can seriously affect a developing embryo or lead to development of cancer. Cancers can arise when proto-oncogenes, which are normally not expressed, become oncogenes that are expressed. On the other hand, cancers can also arise when tumor suppressor genes fail to be adequately expressed.

Although mutations in regulatory genes may represent a major evolutionary force, the mechanisms by which gene expression is regulated also make a major contribution. The proteins in chimpanzees and humans are strikingly similar in amino acid sequence. It's possible that the presence or absence of particular proteins in particular cells may cause this, that new and novel combinations of exons of certain genes may create new proteins that differentiate the two species, or that other changes in the DNA, such as changes in the chromatin structure around certain genes, may account for their phenotypic differences. Throughout the past half century, our knowledge of the mechanisms that regulate gene expression and our understanding of how mutations drive the evolutionary process have blossomed. In Chapter 14, you will see how this knowledge is being harnessed to develop cutting-edge technologies that promise to revolutionize agriculture and medicine.

summary

13.1 Prokaryotic Regulation

Prokaryotes often organize genes that are involved in a common process or pathway into operons in which the genes are coordinately regulated. Gene expression in prokaryotes is usually regulated at the level of transcription. The operon model developed by Jacob and Monod says that a regulator gene codes for a repressor, which sometimes binds to the operator. When it does, RNA polymerase is unable to bind to the promoter, and transcription of the structural genes of the operon cannot take place. However, we now know that operons may be regulated by both activators and repressors.

The *trp* operon is an example of a repressible operon because when tryptophan, the corepressor, is present, it binds to the repressor. The repressor is then able to bind to the operator, and transcription of structural genes does not take place.

The *lac* operon is an example of an inducible operon because when lactose, the inducer, is present, it binds to the repressor. The repressor is unable to bind to the operator, and transcription of structural genes takes place if glucose is absent.

Both the *lac* and *trp* operons exhibit negative control, because a repressor is involved. However, positive control also occurs. There are also examples of positive control. The structural genes in the *lac* operon are not maximally expressed unless glucose is absent and lactose is present. At that time, cAMP attaches to a molecule called CAP, and then CAP binds to a site next to the promoter. Now RNA polymerase is better able to bind to the promoter, and transcription occurs.

13.2 Eukaryotic Regulation

The following levels of control of gene expression are possible in eukaryotes: chromatin structure, transcriptional control, posttranscriptional control, translational control, and posttranslational control.

Chromatin structure helps regulate transcription. Highly condensed heterochromatin is genetically inactive, as exemplified by Barr bodies. Less-condensed euchromatin is genetically active, as exemplified by lampbrush chromosomes in vertebrates.

Regulatory proteins called transcription factors, as well as DNA sequences called enhancers and silencers, play a role in controlling transcription in eukaryotes. Transcription factors bind to the promoter, and transcription activators bind to an enhancer.

Posttranscriptional control is achieved by creating variations in messenger RNA (mRNA) splicing, which may yield multiple mRNA messages from the same gene, and by altering the speed with which a particular mRNA molecule leaves the nucleus.

Translational control affects mRNA translation and the length of time it is translated, primarily by altering the stability of an mRNA. MicroRNAs are a unique example of translational control. Posttranslational control affects whether or not an enzyme is active and how long it is active.

13.3 Regulation Through Gene Mutations

In molecular terms, a gene is a sequence of DNA nucleotide bases, and a genetic mutation is a change in this sequence. Mutations can be spontaneous or due to environmental mutagens such as radiation and organic chemicals. Carcinogens are mutagens that cause cancer.

Point mutations can range in effect, depending on the particular codon change. Sickle cell disease is an example of a point mutation that greatly changes the activity of the affected gene. Frameshift mutations result when a base is added or deleted, and the result is usually a nonfunctional protein. Most cases of cystic fibrosis are due to a frameshift mutation. Nonfunctional proteins can affect the phenotype drastically, as in albinism, which is due to a single faulty enzyme, and androgen insensitivity, which is due to a faulty receptor for testosterone.

Cancer is often due to an accumulation of genetic mutations among genes that code for regulatory proteins. The cell cycle occurs inappropriately when proto-oncogenes become oncogenes and tumor suppressor genes are no longer effective. Mutations that affect transcription factors and other regulators of gene expression are frequent causes of cancer.

understanding the terms

Barr body 238
carcinogen 243
corepressor 235
DNA repair enzyme 244
enhancer 240
epigenetic inheritance 237
euchromatin 239
frameshift mutation 244
gene mutation 243
heterochromatin 238
induced mutation 243
inducer 236
inducible operon 236
microRNA 241
mutagen 243
operator 234
operon 234
point mutation 244
posttranscriptional control 240
posttranslational control 242
promoter 234
regulator gene 235
repressible operon 235
repressor 234
spontaneous mutation 243
structural gene 234
transcription activator 240
transcriptional control 240
transcription factor 240
translational control 241

Match the terms to these definitions:

a. _____ A regulation of gene expression that begins once there is an mRNA transcript.
b. _____ Genes involved in a common function that are located and regulated together.
c. _____ Dark-staining body in the nuclei of female mammals that contains a condensed, inactive X chromosome.
d. _____ Changes in the base sequence of DNA that arise as a result of environmental influences.
e. _____ Environmental agent that causes mutations leading to the development of cancer.

reviewing this chapter

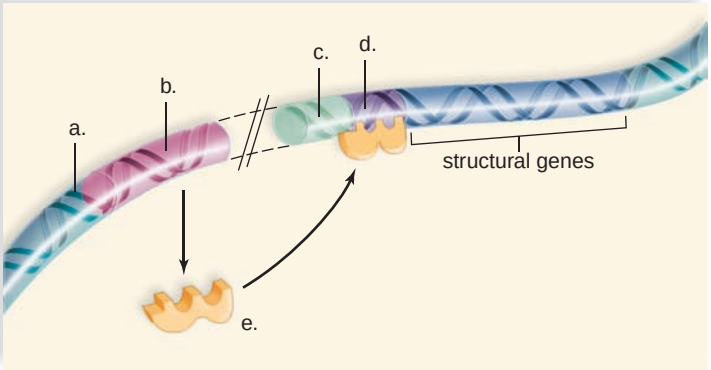
1. Name and state the function of the three components of operons. 234
2. Explain the operation of the *trp* operon, and note why it is considered a repressible operon. 235
3. Explain the operation of the *lac* operon, and note why it is considered an inducible operon. 235–36
4. What are the five levels of genetic regulatory control in eukaryotes? 237
5. Relate heterochromatin and euchromatin to levels of chromatin organization. 238–39
6. With regard to transcriptional control in eukaryotes, explain how Barr bodies show that heterochromatin is genetically inactive. 238–39
7. Explain how lampbrush chromosomes in vertebrates show that euchromatin is genetically active. 239
8. What do transcription factors do in eukaryotic cells? What are enhancers? 240
9. Explain how alternative mRNA processing may create multiple mRNAs from a single gene. 240–41
10. Give examples of translational and posttranslational control in eukaryotes. 241–42
11. Name some causes of mutations. 243
12. What are two major types of mutations, and what effect can they have on protein activity? 243–44
13. Mutations in what types of genes, in particular, can cause cancer? 245

testing yourself

Choose the best answer for each question.

1. Which of the following illustrates negative control?
   a. A repressor that becomes active when bound to a corepressor and inhibits transcription.
   b. A gene that binds a repressor and becomes active.
   c. An activator that becomes active when bound to a coactivator and activates transcription.
   d. A repressor that binds a gene and becomes inactive.
2. In regulation of the *lac* operon, when lactose is present and glucose is absent,
   a. the repressor is able to bind to the operator.
   b. the repressor is unable to bind to the operator.
   c. transcription of structural genes occurs.
   d. transcription of lactose occurs.
   e. Both b and c are correct.
3. In regulation of the *trp* operon, when tryptophan is present,
   a. the repressor is able to bind to the operator.
   b. the repressor is unable to bind to the operator.
   c. transcription of the repressor in inhibited.
   d. transcription of the structural genes, operator, and promoter occurs.
4. In operon models, the function of the promoter is to
   a. code for the repressor protein.
   b. bind with RNA polymerase.
   c. bind to the repressor.
   d. code for the regulator gene.
5. Which of the following statements is/are true regarding operons?
   a. The regulator gene is transcribed with the structural genes.
   b. The structural genes are always transcribed.
   c. All genes are always transcribed.
   d. The regulator gene has its own promoter.
6. Which of the following regulate gene expression in the eukaryotic nucleus?
   a. posttranslational control
   b. transcriptional control
   c. translational control
   d. posttranscriptional control
   e. Both b and d are correct.
7. Which of the following mechanisms may create multiple mRNAs from the same gene?
   a. posttranslational control
   b. alternative mRNA splicing
   c. binding of a transcription factor
   d. chromatin remodeling
   e. miRNAs
8. Translational control of gene expression occurs within the
   a. nucleus.
   b. cytoplasm.
   c. nucleolus.
   d. mitochondria.
9. Alternative mRNA splicing is an example of which type of regulation of gene expression?
   a. transcriptional
   b. posttranscriptional
   c. translational
   d. posttranslational
10. A scientist adds radioactive uridine (label for RNA) to a culture of cells and examines an autoradiograph. Which type of chromatin is apt to show the label?
    a. heterochromatin
    b. euchromatin
    c. the histones, not the DNA
    d. the DNA, not the histones
    e. Both a and d are correct.

11. Barr bodies are
    a. genetically active X chromosomes in males.
    b. genetically inactive X chromosomes in females.
    c. genetically active Y chromosomes in males.
    d. genetically inactive Y chromosomes in females.
12. Which of these might cause a proto-oncogene to become an oncogene?
    a. exposure of the cell to radiation
    b. exposure of the cell to certain chemicals
    c. viral infection of the cell
    d. exposure of the cell to pollutants
    e. All of these are correct.
13. A cell is cancerous. You might find an abnormality in
    a. a proto-oncogene.
    b. a tumor suppressor gene.
    c. regulation of the cell cycle.
    d. tumor cells.
    e. All of these are correct.
14. A tumor suppressor gene
    a. inhibits cell division.
    b. opposes oncogenes.
    c. prevents cancer.
    d. is subject to mutations.
    e. All of these are correct.
15. Label this diagram of an operon.



16. If the DNA codons are CAT CAT CAT, and a guanine base is added at the beginning, which would result?
    a. CAT CAT CAT G
    b. G CAT CAT CAT
    c. GCA TCA TCA T
    d. GC ATC ATC AT
17. A mutation in a DNA molecule involving the replacement of one nucleotide base pair with another is called a(n)
    a. frameshift mutation.
    b. transposon.
    c. deletion mutation.
    d. point mutation.
    e. insertion mutation.
18. Which of these is characteristic of cancer?
    a. It may involve a lack of mutations over a length of time.
    b. It cannot be tied to particular environmental factors.
    c. Apoptosis is one of the first developmental effects.
    d. Mutations in certain types of genes.
    e. It typically develops within a short period of time.
19. Which is not evidence that eukaryotes control transcription?
    a. euchromatin/heterochromatin
    b. existence of transcription factors
    c. lampbrush chromosomes
    d. occurrence of mutations
    e. All of these are correct.

thinking scientifically

1. In patients with chronic myelogenous leukemia, an odd chromosome is seen in all the cancerous cells. A small piece of chromosome 9 is connected to chromosome 22. This 9:22 translocation has been termed the Philadelphia chromosome. How could a translocation cause genetic changes that result in cancer?
2. New findings indicate that mutations outside of genes may cause disease, such as in some cases of Hirschsprung disease and multiple endocrine neoplasia. Explain how such a mutation might alter the expression of a gene.

bioethical issue

Environmental Estrogens and Mutation

You have learned from this chapter that many types of carcinogens, such as those found in cigarette smoke, may alter the base sequence of DNA. However, environmental estrogens are a recently identified type of carcinogen that is generating much attention and concern in recent years. Environmental estrogens are estrogen-like compounds that can disrupt normal endocrine system function in animals by competing with normal sex hormones for receptors, inadvertently activating and inactivating transcription factors and greatly affecting gene expression. They have been linked to increased mutation rates, to deformed genitals in alligators and fish, to promotion of cell division in cultured breast cancer cells, and to inhibition of sperm development in humans.

Environmental estrogens are sometimes found naturally at low concentrations in foods such as soybeans and flax seeds. However, many of these compounds are artificial, originating from chemicals such as polychlorinated biphenyls (PCBs), phthalates (found extensively in many plastics), and atrazine, a compound found in many commercial weed killers. Many people, including scientists at the EPA, contend that these artificial compounds, even at very low doses, are a major threat to the environment, to many animal species, and to human health.

However, some critics contend that the concentrations of these compounds in the soil, air, and water are far below concentrations necessary to cause problems in most animal species, including humans. They also tout studies showing high concentrations of environmental estrogens in many grains, fruits, and vegetables, and that many of these compounds are rendered harmless by the body before they have a chance to cause mutations.

Should known environmental estrogens, such as those found in plastics, herbicides, and insecticides, be closely monitored by the government, and maximal permissible levels set for their emission into the environment? And where should money to fund these regulations be derived? Or, as some critics insist, are we worried about a problem that simply does not exist?

Biology website

The companion website for *Biology* provides a wealth of information organized and integrated by chapter. You will find practice tests, animations, videos, and much more that will complement your learning and understanding of general biology.

http://www.mhhe.com/maderbiology10

14

Biotechnology and Genomics

Biotechnology is used in a myriad of ways today, even to make ice cream more smooth and creamy. An eel-like fish, the ocean pout, produces a natural antifreeze protein that is now made by genetically modified bacteria. The product is readily available to all, even ice-cream manufacturers who want their product to be free of ice crystals. Modern biotechnology has also made it possible for farmers, bioengineers, and medical scientists to alter the genotype and subsequently the phenotype of plants, animals, and humans. Genetically modified crops are resistant to disease and able to grow under stressful conditions. Farm animals grow larger than usual, and humans are supplied with normal genes to make up for ones that do not function as they should.

But others worry that genetically-modified bacteria and plants might harm the environment and that the products produced by altered organisms might not be healthy for humans. Other ethical concerns abound. Is it ethical to give a cat a gene that makes it glow? To what extent would it be proper to improve the human genome? Everyone should be knowledgeable about modern genetics so they can participate in deciding these issues.

A biotechnology product derived from the ocean pout (*Zoarces americanus*) improves the texture of ice cream.

concepts

14.1 DNA Cloning

In biology, **cloning** is the production of genetically identical copies of DNA, cells, or organisms through some asexual means. When an underground stem or root sends up new shoots, the resulting plants are clones of one another. The members of a bacterial colony on a petri dish are clones because they all came from the division of the same original cell. Human identical twins are also considered clones. The first two cells of the embryo separate, and each becomes a complete individual.

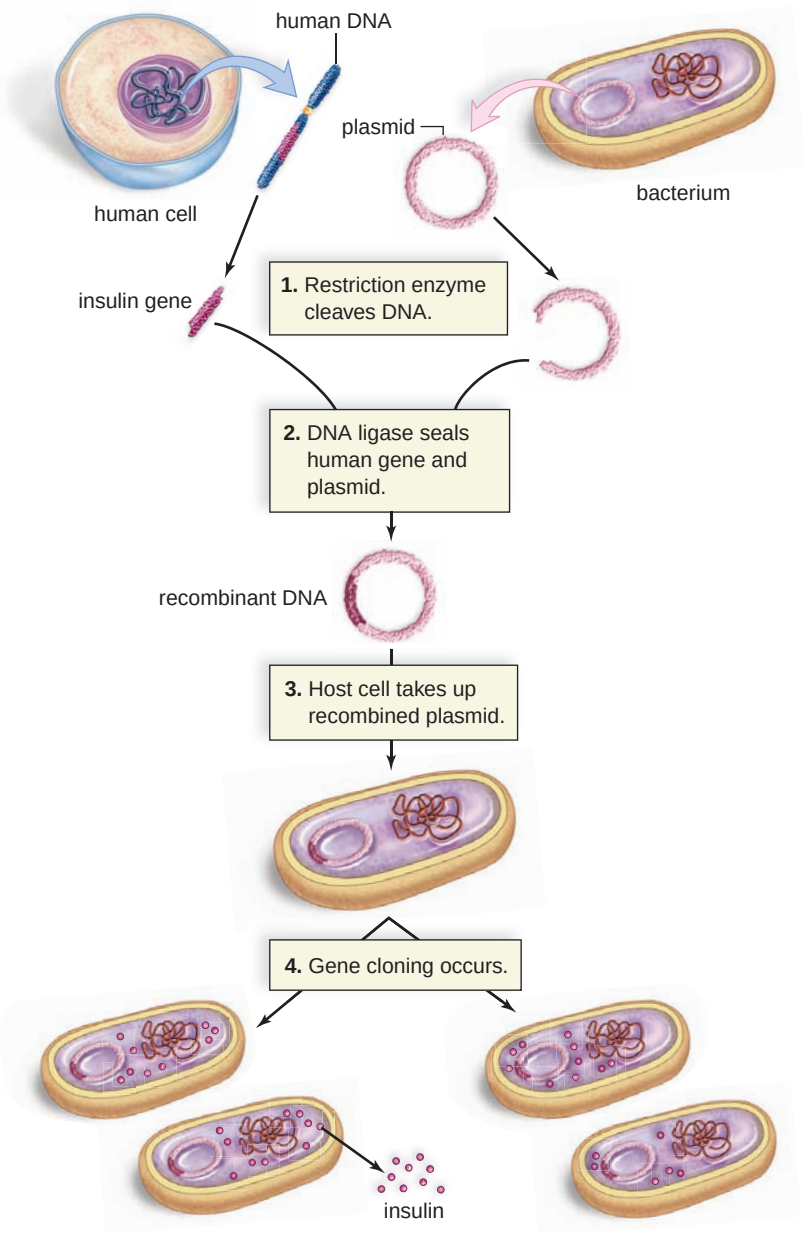


FIGURE 14.1 Cloning a human gene.

Human DNA and plasmid DNA are cleaved by a specific type of restriction enzyme. Then the human DNA, perhaps containing the insulin gene, is spliced into a plasmid by the enzyme DNA ligase. Gene cloning is achieved after a bacterium takes up the plasmid. If the gene functions normally as expected, the product (e.g., insulin) may also be retrieved.

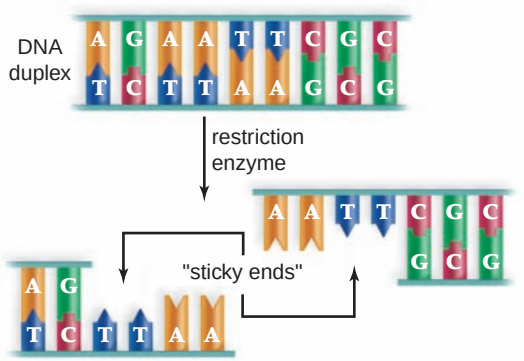
DNA cloning can be done to produce many identical copies of the same gene; that is, for the purpose of **gene cloning.** Scientists clone genes for a number of reasons. They might want to determine the difference in base sequence between a normal gene and a mutated gene. Or they might use the genes to genetically modify organisms in a beneficial way. When cloned genes are used to modify a human, the process is called **gene therapy.** Otherwise, the organisms are called **transgenic organisms** [L. *trans,* across, through; Gk. *genic,* producing]. Transgenic organisms are frequently used today to produce a product desired by humans.

Recombinant DNA (rDNA) technology and the **polymerase chain reaction (PCR)** are two procedures that scientists can use to clone DNA.

Recombinant DNA Technology

Recombinant DNA (rDNA) contains DNA from two or more different sources, such as a human cell and a bacterial cell, as shown in Figure 14.1. To make rDNA, a technician needs a **vector** [L. *vehere,* to carry], by which rDNA will be introduced into a host cell. One common vector is a plasmid. **Plasmids** are small accessory rings of DNA found in bacteria. The ring is not part of the bacterial chromosome and replicates on its own. Plasmids were discovered by investigators studying the bacterium *Escherichia coli (E. coli).*

Two enzymes are needed to introduce foreign DNA into vector DNA: (1) a **restriction enzyme,** which cleaves DNA, and (2) an enzyme called **DNA ligase** [L. *ligo,* bind], which seals DNA into an opening created by the restriction enzyme. Hundreds of restriction enzymes occur naturally in bacteria, where they cut up any viral DNA that enters the cell. They are called restriction enzymes because they *restrict* the growth of viruses, but they also act as molecular scissors to cleave any piece of DNA at a specific site. For example, the restriction enzyme called *Eco*RI always cuts double-stranded DNA at this sequence of bases and in this manner:



Notice that there is now a gap into which a piece of foreign DNA can be placed if it ends in bases complementary to those exposed by the restriction enzyme. The single-stranded, but complementary, ends of the two DNA molecules are called "sticky ends" because they can bind a piece of foreign DNA by complementary base-pairing. Sticky ends facilitate the insertion of foreign DNA into vector DNA as long as both are cleansed by the same restriction enzyme.

Next, genetic engineers use the enzyme DNA ligase to seal the foreign piece of DNA into the vector. DNA splicing is now complete; an rDNA molecule has been prepared (Fig. 14.1). Bacterial cells take up recombinant plasmids, especially if they are treated to make their cell membranes more permeable. Thereafter, as the plasmid replicates, DNA is cloned.

In order for bacteria to express a human gene, the cloned gene has to be accompanied by regulatory regions unique to bacteria. Also, the gene should not contain introns because bacteria don't have introns. However, it is possible to make a human gene that lacks introns. The enzyme called reverse transcriptase can be used to make a DNA copy of human mRNA. The DNA molecule, called **complementary DNA (cDNA),** does not contain introns. Bacteria may then transcribe and translate the cloned cDNA to produce a human protein because the genetic code is the same in humans and bacteria.

The Polymerase Chain Reaction

The polymerase chain reaction (PCR), developed by Kary Mullis in 1985, can create copies of a segment of DNA quickly in a test tube. PCR is very specific—it amplifies (makes copies of) a targeted DNA sequence. The targeted sequence can be less than one part in a million of the total DNA sample!

PCR requires the use of DNA polymerase, the enzyme that carries out DNA replication, and a supply of nucleotides for the new DNA strands. PCR is a chain reaction because the targeted DNA is repeatedly replicated as long as the process continues. The colors in Figure 14.2 distinguish the old strand from the new DNA strand. Notice that the amount of DNA doubles with each replication cycle.

PCR has been in use for years, and now almost every laboratory has automated PCR machines to carry out the procedure. Automation became possible after a temperature-insensitive (thermostable) DNA polymerase was extracted from the bacterium *Thermus aquaticus,* which lives in hot springs. The enzyme can withstand the high temperature used to separate double-stranded DNA; therefore, replication does not have to be interrupted by the need to add more enzyme.

Analyzing DNA

DNA amplified by PCR can be analyzed for various purposes. For example, mitochondrial DNA taken from modern living populations was used to decipher the evolutionary history of human populations. For identification purposes, DNA taken from a corpse burned beyond recognition can be matched to that on the bristles of their toothbrush!

Analysis of DNA following PCR has undergone improvements over the years. At first, the entire genome was treated with restriction enzymes, resulting in a unique collection of different-sized fragments because each person has their own restriction enzyme sites. During a process called gel electrophoresis, the fragments were separated according to their size, and the result was a pattern of distinctive bands that identified the person. Now, **short tandem repeat (STR) profiling** is the method of choice. STRs are the same short sequence of DNA bases that recur several times, as in TCGTCGTCG. STR profiling is advantageous because it doesn't require the use of restriction enzymes. The chromosomal locations for STRs are known and, therefore, it is possible to subject only these locations to PCR and use gel electrophoresis to arrive at a band pattern that is different for each person. The band patterns are different because each person has their own number of repeats at the different locations. The greater the number of STRs at a location, the longer the DNA fragment amplified by PCR. The more STR

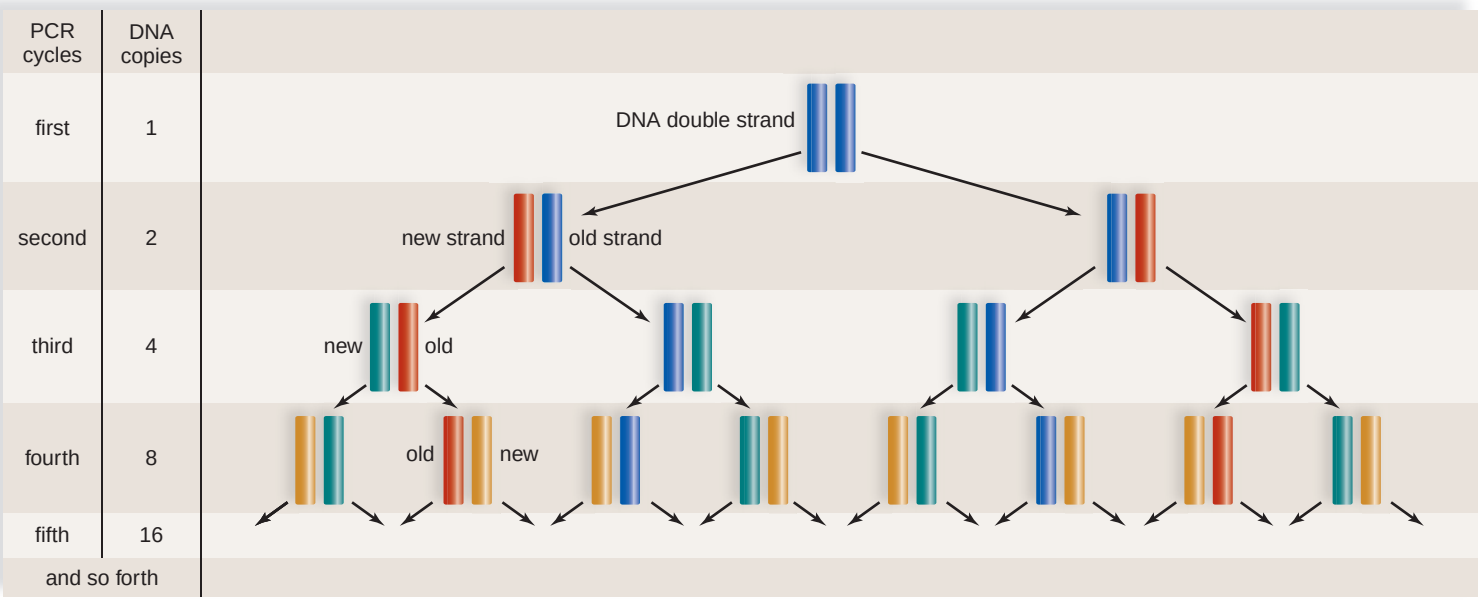
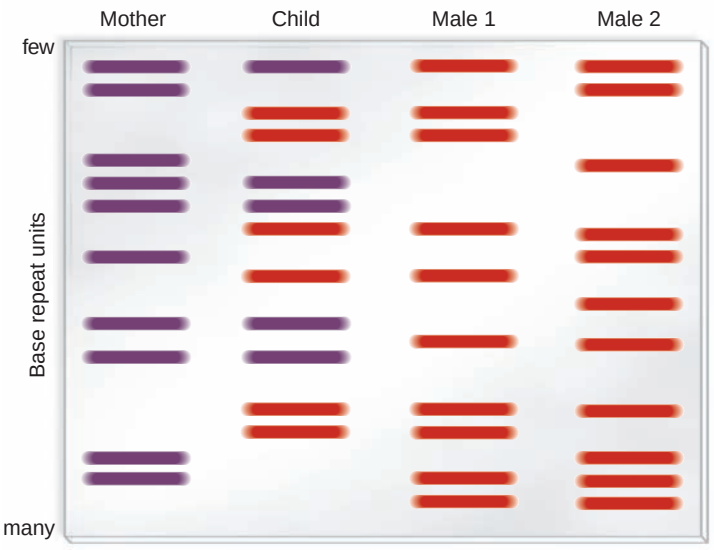


FIGURE 14.2 Polymerase chain reaction (PCR).

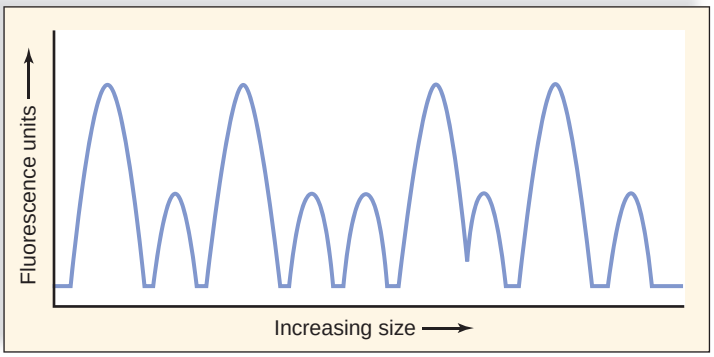
PCR allows the production of many identical copies of DNA in a laboratory setting.

locations employed, the more confident scientists can be of distinctive results for each person (Fig. 14.3*a*). The newest method of doing fingerprints today does away with the need to use gel electrophoresis: The DNA fragments are fluorescently labeled. Using a particular laboratory instrument, a laser excites the fluorescent STRs, and a detector records the amount of emission for each DNA fragment in terms of peaks and valleys. Therefore, the greater the fluorescence, the greater the number of repeats at a location. The printout, such as the one shown in Figure 14.3*b*, is the DNA fingerprint, and each person has their own particular printout.

Applications of PCR are limited only by our imaginations. When the DNA matches that of a virus or mutated gene, it is known that a viral infection, genetic disorder, or cancer is present. DNA fingerprinted from blood or tissues at a crime scene has been successfully used in convicting criminals. DNA fingerprinting through STR profiling was extensively used to identify the victims of the September 11, 2001 terrorist attacks in the United States. Relatives can be found, paternity suits can be settled (Fig. 14.3*a*), genetic disorders can be detected, and illegally poached ivory and whale meat can be recognized using this technology. PCR has also shed new light on evolutionary studies by comparing extracted DNA from certain fossils with that of living organisms.



a. DNA Band patterns



b. Automated DNA fingerprinting

FIGURE 14.3 The use of DNA fingerprints to establish paternity.

a. In this method, DNA fragments containing STRs are separated by gel electrophoresis. Male 1 is the father. **b.** Each person's fingerprint pattern (only one is shown) can also be printed out by a machine that detects fluorescence.

Check Your Progress 14.1

1. Describe the process of creating an rDNA molecule.
2. How can DNA fingerprinting be used to analyze DNA molecules?

14.2 Biotechnology Products

Today, transgenic bacteria, plants, and animals are often called **genetically modified organisms (GMOs),** and the products they produce are called **biotechnology products.**

Genetically Modified Bacteria

Recombinant DNA technology is used to produce transgenic bacteria, which are grown in huge vats called bioreactors. The gene product is usually collected from the medium. Biotechnology products produced by bacteria include insulin, clotting factor VIII, human growth hormone, t-PA (tissue plasminogen activator), and hepatitis B vaccine. Transgenic bacteria have many other uses as well. Some have been produced to promote the health of plants. For example, bacteria that normally live on plants and encourage the formation of ice crystals have been changed from frost-plus to frost-minus bacteria. As a result, new crops such as frost-resistant strawberries are being developed. Also, a bacterium that normally colonizes the roots of corn plants has now been endowed with genes (from another bacterium) that code for an insect toxin. The toxin protects the roots from insects.

Bacteria can be selected for their ability to degrade a particular substance, and this ability can then be enhanced by bioengineering. For instance, naturally occurring bacteria that eat oil can be genetically engineered to do an even better job of cleaning up beaches after oil spills (Fig. 14.4). Bacteria can also remove sulfur from coal before it is burned and help clean up toxic waste dumps. One such strain was given genes that allowed it to clean up levels of toxins that would have killed other strains. Further, these bacteria were given "suicide" genes that caused them to self-destruct when the job had been accomplished.

Organic chemicals are often synthesized by having catalysts act on precursor molecules or by using bacteria to carry out the synthesis. Today, it is possible to go one step further and manipulate the genes that code for these enzymes. For instance, biochemists discovered a strain of bacteria that is especially good at producing phenylalanine, an organic chemical needed to make aspartame, the dipeptide sweetener better known as NutraSweet. They isolated, altered, and formed a vector for the appropriate genes so that various bacteria could be genetically engineered to produce phenylalanine.

FIGURE 14.4
Genetically modified bacteria.

Bacteria capable of decomposing oil have been engineered and patented. In the inset, the flask toward the rear contains oil and no bacteria; the flask toward the front contains the bacteria and is almost clear of oil.

Genetically Modified Plants

Techniques have been developed to introduce foreign genes into immature plant embryos or into plant cells called *protoplasts* that have had the cell wall removed. It is possible to treat protoplasts with an electric current while they are suspended in a liquid containing foreign DNA. The electric current makes tiny, self-sealing holes in the plasma membrane through which genetic material can enter. Protoplasts go on to develop into mature plants. One altered plant known as the pomato is the result of these technologies. This plant produces potatoes belowground and tomatoes aboveground.

Foreign genes transferred to cotton, corn, and potato strains have made these plants resistant to pests because their cells now produce an insect toxin. Similarly, soybeans have been made resistant to a common herbicide. Some corn and cotton plants are both pest- and herbicide-resistant. These and other genetically modified crops that are expected to have increased yield are now sold commercially.

Like bacteria, plants are also being engineered to produce human proteins, such as hormones, clotting factors, and antibodies, in their seeds. One type of antibody made by corn can deliver radioisotopes to tumor cells, and another made by soybeans can be used to treat genital herpes.

Genetically Modified Animals

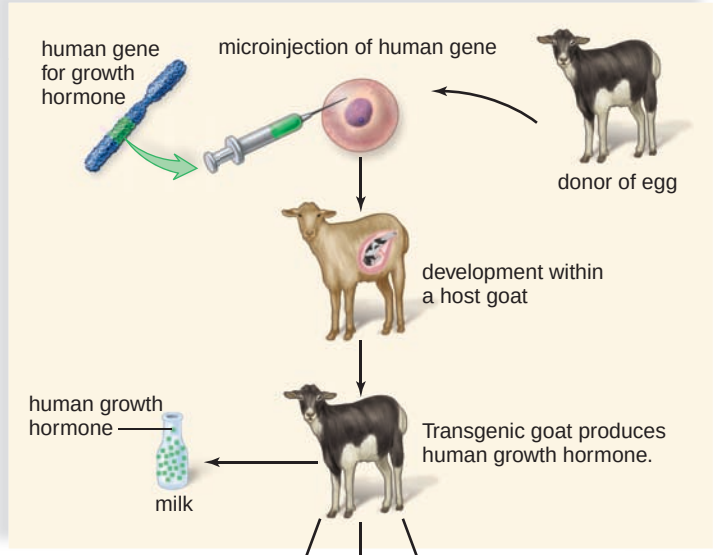
Techniques have been developed to insert genes into the eggs of animals. It is possible to microinject foreign genes into eggs by hand, but another method uses vortex mixing. The eggs are placed in an agitator with DNA and silicon-carbide needles, and the needles make tiny holes through which the DNA can enter. When these eggs are fertilized, the resulting offspring are transgenic animals. Using this technique, many types of animal eggs have taken up the gene for bovine growth hormone (bGH). The procedure has been used to produce larger fishes, cows, pigs, rabbits, and sheep.

Gene pharming, the use of transgenic farm animals to produce pharmaceuticals, is being pursued by a number of firms. Genes that code for therapeutic and diagnostic proteins are incorporated into an animal's DNA, and the proteins appear in the animal's milk. Plans are under way to produce drugs for the treatment of cystic fibrosis, cancer, blood diseases, and other disorders by this method. Figure 14.5*a* outlines the procedure for producing transgenic mammals: DNA containing the gene of interest is injected into donor eggs. Following in vitro fertilization, the zygotes are placed in host females, where they develop. After female offspring mature, the product is secreted in their milk.

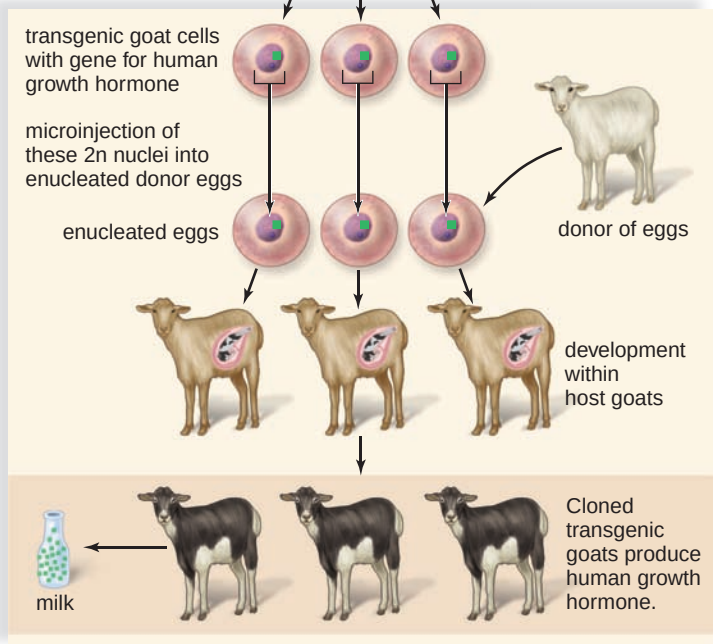
Cloning Transgenic Animals

For many years, it was believed that adult vertebrate animals could not be cloned because cloning requires that all the genes of an adult cell be turned on if development is to proceed normally. This had long been thought impossible.

In 1997, however, Scottish scientists announced that they had produced a cloned sheep called Dolly. Since then, calves and goats have also been cloned, as described in Figure 14.5*b*. After enucleated eggs from a donor are microinjected with 2n



a.



b.

FIGURE 14.5 Transgenic mammals produce a product.

a. A bioengineered egg develops in a host to create a transgenic goat that produces a biotechnology product in its milk. **b.** Nuclei from the transgenic goat are transferred into donor eggs, which develop into cloned transgenic goats.

nuclei from the same transgenic animal, they are coaxed to begin development in vitro. Development continues in host females until the clones are born. The female clones have the same product in their milk as the donor of the eggs. Now that scientists have a way to clone animals, this procedure will undoubtedly be used routinely to procure biotechnology products. However, animal cloning is a difficult process with a low success rate. The vast majority of cloning attempts are unsuccessful, resulting in the early death of the clone.

Applications of Transgenic Animals

Researchers are using transgenic mice for various research projects. Figure 14.6 shows how this technology has demonstrated that a section of DNA called *SRY* (sex determining region of the Y chromosome) produces a male animal. The *SRY* gene was cloned, and then one copy was injected into one-celled mouse embryos. Injected embryos developed into males, but any that were not injected developed into females.

Eliminating a gene is another way to study a gene's function. A *knockout mouse* has had both alleles of a gene removed or made nonfunctional. For example, scientists have constructed a knockout mouse lacking the *CFTR* gene, the same gene mutated in cystic fibrosis patients. The mutant mouse has a phenotype similar to a human with cystic fibrosis and can be used to test new drugs for the treatment of the disease.

Xenotransplantation is the use of animal organs, instead of human organs, in transplant patients. Scientists have chosen to work with pigs because they are prolific and have long been raised as a meat source. Pigs will be genetically modified to make their organs less likely to be rejected by the human body. The hope is that one day a pig organ will be as easily accepted by the human body as a blood transfusion from a person with the same blood type.

Check Your Progress **14.2**

1. What difficulties are there in creating transgenic animals versus transgenic bacteria?
2. How do transgenic animals differ from cloned animals?

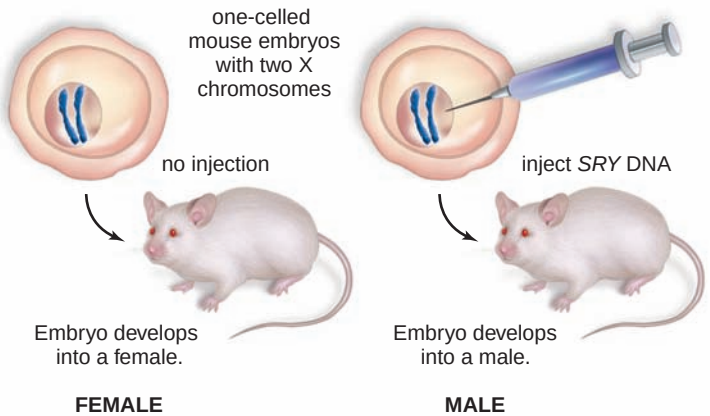


FIGURE 14.6 **Experimental use of mice.**

Bioengineered mice showed that maleness is due to *SRY* DNA.

14.3 Gene Therapy

The manipulation of an organism's genes can be extended to humans in a process called gene therapy. Gene therapy is an accepted therapy for the treatment of a disorder. Gene therapy has been used to cure inborn errors of metabolism, as well as to treat more generalized disorders such as cardiovascular disease and cancer. Figure 14.7 shows regions of the body that have received copies of normal genes by various methods of gene transfer. Viruses genetically modified to be safe can be used to ferry a normal gene into the body, and so can liposomes, which are microscopic globules of lipids specially prepared to enclose the normal gene. On the other hand, sometimes the gene is injected directly into a particular region of the body. Below we give examples of **ex vivo gene therapy** (the gene is inserted into cells that have been removed and then returned to the body) and **in vivo gene therapy** (the gene is delivered directly into the body).

Ex Vivo Gene Therapy

Children who have SCID (severe combined immunodeficiency) lack the enzyme ADA (adenosine deaminase), which is involved in the maturation of immune cells. Therefore, these children are prone to constant infections and may die without treatment. To carry out gene therapy, bone marrow stem cells are removed from the bone marrow of the patient and infected with a virus that carries a normal gene for the enzyme into their DNA. Then the cells are returned to the patient, where it is hoped they will divide to produce more blood cells with the same genes.

Familial hypercholesterolemia is a condition that develops when liver cells lack a receptor protein for removing cholesterol from the blood. The high levels of blood cholesterol make the patient subject to fatal heart attacks at a young age. A small portion of the liver is surgically excised and then infected with a virus containing a normal gene for the receptor before being returned to the patient. Patients are expected to experience lowered serum cholesterol levels following this procedure.

In Vivo Gene Therapy

Cystic fibrosis patients lack a gene that codes for a transmembrane carrier of the chloride ion. They often suffer from numerous and potentially deadly infections of the respiratory tract. In gene therapy trials, the gene needed to cure cystic fibrosis is sprayed into the nose or delivered to the lower respiratory tract by adenoviruses or by the use of a liposome. So far, this treatment has resulted in limited success.

It has been known for some time that VEGF (vascular endothelial growth factor) can cause the growth of new blood vessels. The gene that codes for this growth factor can be injected alone or within a virus into the heart to stimulate branching of coronary blood vessels to increase circulation to the heart. Patients report that they have less chest pain and can run longer on a treadmill.

In cancer patients, genes are being used to make healthy cells more tolerant of chemotherapy and to make tumors more vulnerable to chemotherapy. The gene *p53* brings about apoptosis, and there is much interest in introducing it into cancer cells that no longer have the gene and in that way killing them off.

Check Your Progress 14.3

1. What methods are being used to introduce genes into human beings for gene therapy?
2. Give an example of ex vivo and of in vivo gene therapy.

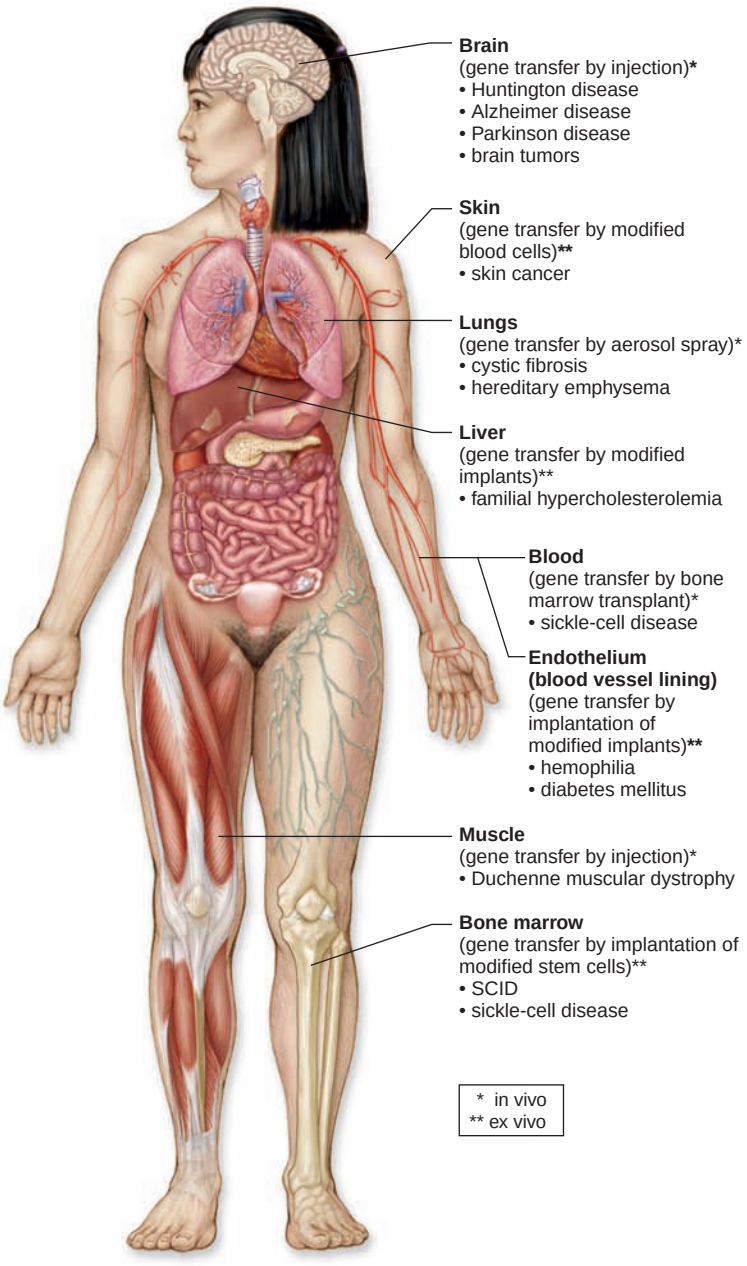


FIGURE 14.7 Gene therapy.

Sites of ex vivo and in vivo gene therapy to cure the conditions noted.

14.4 Genomics

In the previous century, researchers discovered the structure of DNA, how DNA replicates, and how DNA and RNA are involved in the process of protein synthesis. Genetics in the twenty-first century concerns **genomics,** the study of genomes—our complete genetic makeup and also that of other organisms. Knowing the sequence of bases in genomes is the first step, and thereafter we want to understand the function of our genes and their introns and the intergenic sequences (see Figure 14.8). The enormity of the task can be appreciated by knowing that there are approximately 6 billion base pairs in the 2n human genome. Many other organisms have a larger number of protein-coding genes but less noncoding regions compared to the human genome.

Sequencing the Genome

We now know the order of the base pairs in our genome. This feat, which has been likened to arriving at the periodic table of the elements in chemistry, was accomplished by the **Human Genome Project (HGP),** a 13-year effort that involved both university and private laboratories around the world. How did they do it? First, investigators developed a laboratory procedure that would allow them to decipher a short sequence of base pairs, and then instruments became available that could carry out sequencing automatically. Over the 13-year span, DNA sequencers were constantly improved, and now modern instruments can automatically analyze up to 2 million base pairs of DNA in a 24-hour period. Where did all this DNA come from? Sperm DNA was the material of choice because it has a much higher ratio of DNA to protein than other types of cells. (Recall that sperm do provide both X and Y chromosomes.) However, white cells from the blood of female donors were also used in order to include female-originated samples. The male and female donors were of European, African, American (both North and South), and Asian ancestry.

Many small regions of DNA that vary among individuals (polymorphisms) were identified during the HGP. Most of these are *single nucleotide polymorphisms (SNPs)* (a difference of only one nucleotide). Many SNPs have no effect; others may contribute to enzymatic differences affecting the phenotype. It's possible that certain SNP patterns change an individual's susceptibility to disease and alter their response to medical treatments (see p. 284).

Determining that humans have 20,500 genes required a number of techniques, many of which relied on identifying RNAs in cells and then working backward to find the DNA that can pair with that RNA. **Structural genomics**—knowing the sequence of the bases and how many genes we have—is now being followed by functional genomics. Most of the known 20,500 human genes are expected to code for proteins. However, much of the human genome was formerly described as "junk" because it does not specify the order of amino acids in a polypeptide. However, it is possible for RNA molecules to have a regulatory effect in cells, as discussed in the next section.

Eukaryotic Gene Structure

Historically, genes were defined as discrete units of heredity that corresponded to a locus on a chromosome (see Fig. 11.4). While prokaryotes typically possess a single circular chromosome with genes that are packed together very closely, eukaryotic chromosomes are much more complex. The genes are seemingly randomly distributed along the length of a chromosome and are fragmented into exons, with intervening sequences called introns scattered throughout the length of the gene (Fig. 14.8). In general, more complex organisms have more complex genes with more and larger introns. In humans, 95% or more of the average protein-coding gene is introns. Once a gene is transcribed, the introns must be removed and the exons joined together to form a functional mRNA transcript (see Fig. 12.13).

Once regarded as merely intervening sequences, introns are now attracting attention as regulators of gene expression. The presence of introns allows exons to be put together in various sequences so that different mRNAs and proteins can result from a single gene. It could also be that introns function to regulate gene expression and help determine which genes are to be expressed and how they are to be spliced. In fact, entire genes have been found embedded within the introns of other genes.

Intergenic Sequences

DNA sequences occur between genes and are referred to as **intergenic sequences.** In general, as the complexity of an organism increases, so does the proportion of its non-protein-coding DNA sequences. Intergenic sequences are now known to comprise the vast majority of human chromosomes, and protein-coding genes represent only about 1.5% of our total DNA. The remainder of this DNA, once dismissed as "junk DNA," is now thought to serve many important functions, and has piqued the curiosity of many investigators. There are several basic types of intergenic sequences found in the human genome, including (1) repetitive elements, (2) transposons, and (3) unknown sequences. The majority of intergenic sequences are the unknown sequences.

Repetitive DNA Elements

Repetitive DNA elements occur when the same sequence of two or more nucleotides are repeated many times along the length of one or more chromosomes. Repetitive elements are very common—comprising nearly half of the human genome—and, therefore, many scientists believe that their true significance has yet to be discovered. Although many scientists still dismiss them as having no function, others point out that the centromeres and telomeres of chromosomes are composed of repetitive elements and, therefore, repetitive DNA elements may not be as useless as once thought. For one thing, those of the centromere could possibly help with segregating the chromosomes during cell division.

Repetitive DNA elements occur as tandem repeats and interspersed repeats. **Tandem repeat** means that the repeated sequences are next to each other on the chromosome.

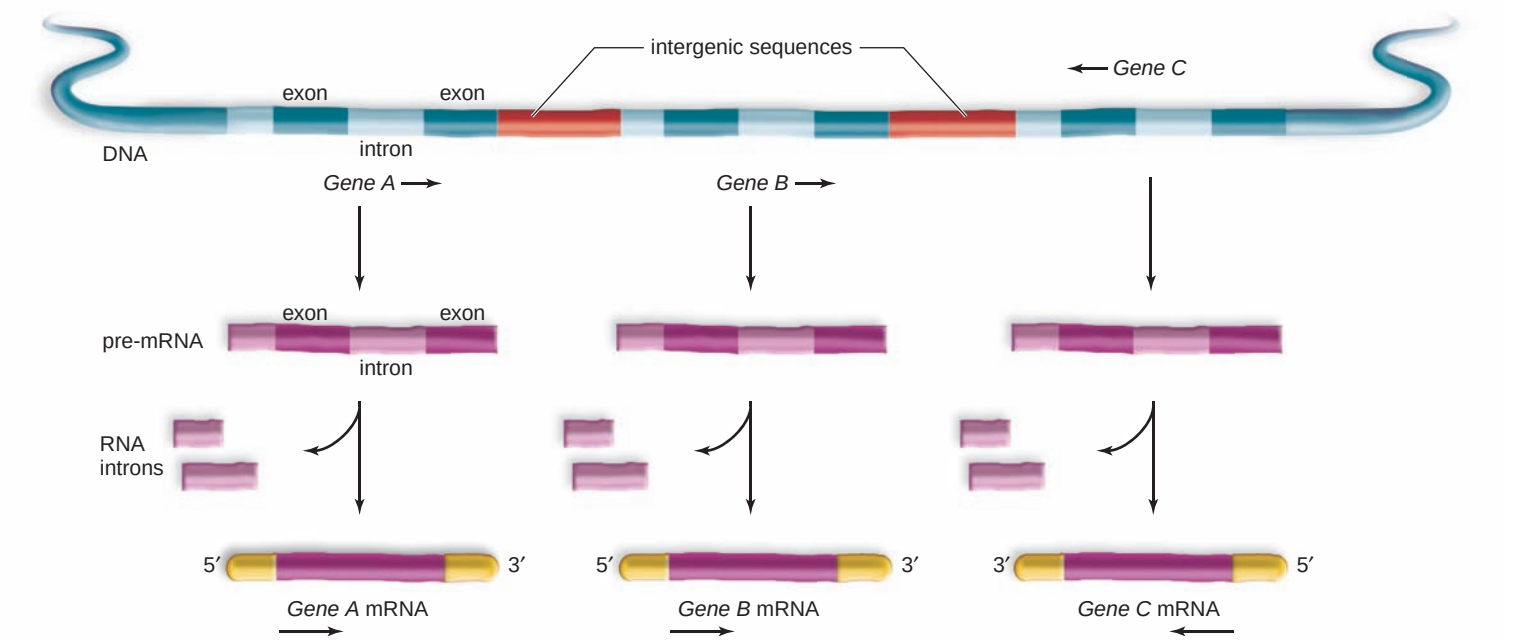


FIGURE 14.8 Chromosomal DNA.

Protein-coding DNA sequences are spread throughout eukaryotic chromosomes, often with very long intergenic sequences between them. Which strand of DNA is used as a template strand can vary, and protein-coding sequences may sometimes even be found within the introns of other genes. The functions of non-protein-coding DNA present in introns and intergenic sequences are being investigated extensively today.

Tandem repeats are often referred to as satellite DNA, because they have a different density than the rest of the DNA within the chromosome. The number and types of tandem repeats may vary significantly from one individual to another, making them invaluable as indicators of heritage. One type of tandem repeat sequence, referred to as *short tandem repeats,* or STRs, has become a standard method in forensic science for identifying one individual from another and for determining familial relationships (see page 251).

The second type of repetitive DNA element is called an **interspersed repeat,** meaning that the repetitions may be placed intermittently along a single chromosome, or across multiple chromosomes. For example, a repetitive DNA element, known as the *Alu* sequence, is interspersed every 5,000 base pairs in human DNA and comprises nearly 5–6% of total human DNA. Because of their common occurrence, interspersed repeats are thought to play a role in the evolution of new genes.

Transposons

Transposons are specific DNA sequences that have the remarkable ability to move within and between chromosomes. Their movement to a new location sometimes alters neighboring genes, particularly decreasing their expression. In other words, a transposon sometimes acts like a regulator gene. The movement of transposons throughout the genome is thought to be a driving force in the evolution of living things. In fact, many scientists now think that many repetitive DNA elements were originally derived from transposons.

Although Barbara McClintock first described these "movable elements" in corn over 50 years ago, it took time for the scientific community to fully appreciate this revolutionary idea. In fact, their significance was only realized within the past few decades. So-called jumping genes now have been discovered in bacteria, fruit flies, humans, and many other organisms. McClintock received a Nobel Prize in 1983 for her discovery of transposons and for her pioneering work in genetics.

Unknown Sequences

While genes comprise a scant 1.5% of the human genome and repetitive DNA elements make up about 44%, the function of the remaining half, called unknown sequences, remains a mystery. Even though this DNA does not appear to contain any protein-coding genes, it has been highly conserved through evolution. In the many millions of years that separates humans from mice, large tracts of this mysterious DNA have remained almost unchanged. But if it has no relevant function, then why has it been so meticulously maintained?

Recently, scientists observed that between 74% and 93% of the genome is transcribed into RNA, including many of these unknown sequences. Thus, what was once thought to be a vast junk DNA wasteland may be much more important than once thought and may play active roles in the cell. Small-sized RNAs may be able to carry out regulatory functions more easily than proteins at times. Therefore, a heretofore-overlooked RNA signaling network may be what allows humans, for example, to achieve structural complexity far beyond anything seen in the unicellular world. Together, these findings have revealed a much more complex, dynamic genome than was envisioned merely a few decades ago.

What Is a Gene?

Perhaps the modern definition of a gene should take the emphasis away from the chromosome and place it on the results of transcription. Previously, molecular genetics considered a gene to be a nucleic acid sequence that codes for the sequence of amino acids in a protein. In contrast to this definition, we have known for some time all three types of RNA are transcribed from DNA and that these RNAs are useful products. We also know that protein-coding regions can be interrupted by regions that do not code for a protein but do produce RNAs with various functions. In recognition of these new findings, what about using this definition suggested by Mark Gerstein and associates in 2007: "A gene is a genomic sequence (either DNA or RNA) directly encoding functional products, either RNA or protein."[1]

This definition merely expands on the central dogma of genetics and recognizes that a gene product need not be a protein, and a gene need not be a particular locus on a chromosome. The DNA sequence that results in a gene product can be split and be present on one or several chromosomes. Also, any DNA sequence can result in one or more products.

This definition recognizes that some prokaryotes have RNA genes. In other words, the genetic material need not be DNA. Again, we can view this as a simple expansion of the central dogma of genetics.

The definition does not spell out what is meant by functional product. It would seem, then, that sequences of DNA resulting in regulatory RNAs or proteins could be considered genes.

The definition does expand on what is meant by "coding." Coding does not necessarily mean a DNA sequence that codes for a sequence of amino acids. Coding simply means a sequence of DNA bases that are transcribed. The gene product can be RNA molecules, or it can be a protein.

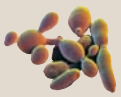
Check Your Progress 14.4A

1. How does a tandem repeat differ from an interspersed repeat?
2. Why is a new definition of a gene required?

[1] Mark B. Gerstein et al., *What is a gene, post-encode? History and Updated Definition,* Cold Spring Harbor Laboratory Press, 2007.

TABLE 14.1

Comparison of Sequenced Genomes

Organism	*Homo sapiens* (human) 	*Mus musculus* (mouse) 	*Drosophila melanogaster* (fruit fly) 	*Arabidopsis thaliana* (flowering plant) 	*Caenorhabditis elegans* (roundworm) 	*Saccharomyces cerevisiae* (yeast) 
Estimated Size	2,900 million bases	2,500 million bases	180 million bases	125 million bases	97 million bases	12 million bases
Estimated Number of Genes	~20,500	~30,000	13,600	25,500	19,100	6,300
Chromosome Number	46	40	8	10	12	32

Functional and Comparative Genomics

Since we now know the structure of our genome, the emphasis today is on functional genomics and also on comparative genomics. The aim of **functional genomics** is to understand the exact role of the genome in cells or organisms. To that end, a new technology called **DNA microarrays** can be used to monitor the expression of thousands of genes simultaneously. In other words, the use of a microarray can tell you what genes are turned on in a specific cell or organism at a particular time and under what particular environmental circumstances. The Science Focus on page 259 discusses the importance of DNA microarrays, which are also known as DNA chips, or genome chips. DNA microarrays contain microscopic amounts of known DNA sequences fixed onto a small glass slide or silicon chip in known locations (see Fig. 14A). When mRNA molecules of a cell or organism bind through complementary base pairing with the various DNA sequences, then that gene is active in the cell. As also discussed on page 259, DNA microarrays are available that rapidly identify all the various mutations in the genome of an individual. This is called the person's **genetic profile.** The genetic profile can indicate if any genetic illnesses are likely and what type of drug therapy for an illness might be most appropriate for that individual.

Aside from the protein-coding regions, researchers also want to know how SNPs and non-protein-coding regions, including repeats, affect which proteins are active in cells. As already discussed at length in Chapter 12, much research is now devoted to knowing the function of DNA regions that do not code for proteins.

The aim of **comparative genomics** is to compare the human genome to the genome of other organisms, such as those listed in Table 14.1. Surprisingly, perhaps, functional genomics has also been advanced by sequencing the genome of these organisms called model organisms (Table 14.1). Model organisms are used in genetic analysis because they have many genetic mechanisms and cellular pathways in common with each other and with humans. As described on page 254, much has been learned by genetically modifying mice. However, other model organisms can sometimes be used instead. Scientists inserted a human gene associated with early-onset Parkinson's disease into *Drosophila melanogaster,* and the flies showed symptoms similar to those seen in humans with the disorder. This suggested that we might be able to use these organisms instead of mice to test therapies for Parkinsons.

Comparative genomics also offers a way to study changes in a genome over time because the model organisms have a shorter generation time than humans. Comparing genomes will also help us understand the evolutionary relationships between organisms. One surprising discovery is that the genomes of all vertebrates are similar. Researchers were not surprised to find that the genes of chimpanzees and humans are 98% alike, but they did not expect to find that our sequence is also 85% like that of a mouse. Genomic comparisons will likely reveal evolutionary relationships between organisms never previously considered.

Proteomics

The entire collection of a species' proteins is the **proteome.** At first, it may be surprising to learn that the proteome is larger than the genome until we consider all the many regulatory mechanisms, such as alternative pre-mRNA splicing, that increase the number of possible proteins in an organism.

Proteomics is the study of the structure, function, and interaction of cellular proteins. Specific regulatory mechanisms differ between cells, and these differences account for the specialization of cells. One goal of proteomics is to identify and determine the function of the proteins within a particular cell type. Each cell produces thousands of different proteins that can vary between cells and within the same cell, depending on circulations. Therefore, the goal of proteomics is an overwhelming endeavor. Microarray

science focus

DNA Microarray Technology

With advances in robotic technology, it is now possible to spot all 20,500 known human genes, or even the entire human genome, onto a single microarray (Fig. 14A). The mRNA from the organism or the cell to be tested is labeled with a fluorescent dye and added to the chip. When the mRNAs bind to the microarray, a fluorescent pattern results that is recorded by a computer. Now the investigator knows what DNA is active in that cell or organism. A researcher can use this method to determine the difference in gene expression between two different cell types, such as between liver cells and muscle cells.

A mutation microarray, the most common type, can be used to generate a person's genetic profile. The microarray contains hundreds to thousands of known disease-associated mutant gene alleles. Genomic DNA from the individual to be tested is labeled with a fluorescent dye, and then added to the microarray. The spots on the microarray fluoresce if the individual's DNA binds to the mutant genes on the chip, indicating that the individual may have a particular disorder or is at risk for developing it later in life. This technique can generate a genetic profile much more quickly and inexpensively than older methods involving DNA sequencing.

Diseased Tissues

DNA microarrays also promise to hasten the identification of genes associated with diseased tissues. In the first instance, mRNA derived from diseased tissue and normal tissue is labeled with different fluorescent dyes. The normal tissue serves as a control. The investigator applies the mRNA from both normal and abnormal tissue to the microarray. The relative intensities of fluorescence from a spot on the microarray indicate the amount of mRNA originating from that gene in the diseased tissue relative to the normal tissue. If a gene is activated in the disease, more copies of mRNA will bind to the microarray than from the control tissue, and the spot will appear more red than green.

Genomic microarrays are also used to identify links between disease and chromosomal variations. In this instance, the chip contains genomic DNA that is cut into fragments. Each spot on the microarray corresponds to a known chromosomal location. Labeled genomic DNA from diseased and control tissues bind to the DNA on the chip, and the relative fluorescence from both dyes is determined. If the number of copies of any particular target DNA has increased, more sample DNA will bind to that spot on the microarray relative to the control DNA, and a difference in fluorescence of the two dyes will be detected. Researchers are currently using this technique to identify disease-associated copy number variations, such as those discussed in the Science Focus on page 260.

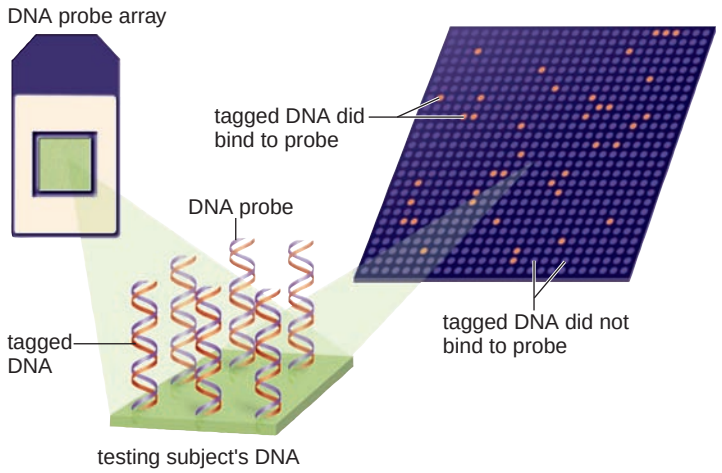


FIGURE 14A DNA microarray technology.
A DNA microarray contains many microscopic samples of DNA bound to known locations on a silicon chip. A fluorescently labeled mRNA from a tissue or organism binds to the DNA on the chip by complementary base-pairing. The fluorescent spots indicate that binding has occurred and that the gene functions in that cell.

technology can assist with this project and so can today's supercomputers.

Computer modeling of the three-dimensional shape of cellular proteins is also an important part of proteomics. If the primary structure of a protein is known, it should be possible to predict its final three-dimensional shape, and even the effects of DNA mutations on the protein's shape and function.

The study of protein function is viewed as essential to the discovery of new and better drugs. Also, it may be possible one day, to correlate drug treatment to the particular proteome of the individual to increase efficiency and decrease side effects. Proteomics will be a critical field of endeavor for many years to come.

Bioinformatics

Bioinformatics is the application of computer technologies, specially developed software, and statistical techniques to the study of biological information, particularly databases that contain much genomic and proteomic information (Fig. 14.9). The new data produced by structural genomics and proteomics have produced raw data that is stored in databases that are readily available to research scientists. It is called raw data because, as yet, it has little meaning. Functional genomics and proteomics are dependent on computer analysis to find significant patterns in the raw data. For example, BLAST, which stands for *b*asic *l*ocal *a*lignment search *t*ool, is a computer program that can identify homologous genes among the genomic sequences of model

organisms. **Homologous genes** are genes that code for the same proteins, although the base sequence may be slightly different. Finding these differences can help trace the history of evolution among a group of organisms.

Bioinformatics also has various applications in human genetics. For example, researchers found the function of the protein that causes cystic fibrosis by using the computer to search for genes in model organisms that have the same sequence. Because they knew the function of this same gene in model organisms, they could deduce the function in humans. This was a necessary step toward possibly developing specific treatments for cystic fibrosis. The human genome has 3 billion known base pairs, and without the computer it would be almost impossible to make sense of these data. For example, it is now known that an individual's genome often contains multiple copies of a gene. But individuals may differ as to the number of copies—called copy number variations, as discussed in the Science Focus below. Now it seems that the number of copies in a genome can be associated with specific diseases. The computer can help make correlations between genomic differences among large numbers of people and disease.

It is safe to say that without bioinformatics, our progress in determining the function of DNA sequences; in comparing our genome to model organisms; in knowing how genes and proteins interact in cells; and so forth, would be extremely slow. Instead, with the help of bioinformatics, progress should proceed rapidly in these and other areas.

FIGURE 14.9 Bioinformatics.

New computer programs are being developed to make sense out of the raw data generated by genomics and proteomics. Bioinformatics allows researchers to study both functional and comparative genomics in a meaningful way.

Check Your Progress 14.4B

1. How is the information learned through the Human Genome Project being used to improve human health?
2. What kind of information can be learned through proteomics?
3. How is comparative genomics being used to divulge evolutionary relationships between organisms?

science focus

Copy Number Variations

Geneticists have long been aware of large chromosomal duplications, deletions, and rearrangements detectable microscopically (see Fig. 10.13*a,b*). However, scientists have recently become aware of small duplications and deletions referred to as copy number variations (CNVs). CNVs occur when *genes* have changed their number.

The change may arise from so-called fork stalling and template switching. DNA damage or some other difficulty may cause the replication fork to stall. In order to continue, the replication machinery can switch to nearby chromosomal material of the same sequence. The replication fork is soon transferred back to the normal template, but the end result is extra or missing copies of small DNA segments. The fact that repetitive elements facilitate template switching suggests a new function for such sequences in our genome.

Some CNVs have known links to disease. Research shows that individuals with fewer copies of the *CCL3L1* gene are more susceptible to HIV infection than those with more copies. Lupus is much more common among people with fewer copies of the complement component *C4* gene. But more surprising was a recent study that suggested at least some cases of autism can be linked to CNVs. The scientists who published the study examined the total chromosomal content of 1,441 autistic children and compared their DNA to more than 2,800 normal individuals. They found that in autistic children, a 25-gene region of chromosome 16 was missing. Furthermore, analysis of other DNA databases revealed the same result: Approximately 1% of autism cases could be directly linked to the same deletion.

CNVs are also emerging as a possible driving force in evolution. A recent study utilized DNA microarrays to examine the chromosomal structure of 47 individuals from many ethnic backgrounds, and found 119 regions where copy number variations existed. More surprising, none of the CNVs were found exclusively in one ethnic group, suggesting that these variants existed well before the human population spread across the Earth. Perhaps they contributed to the phenotypic variations that developed thereafter. Furthermore, many scientists are suggesting that it be advantageous for a species to have multiple copies of genes—if one or both normal copies of an allele fail to function properly, having a third allele available might be advantageous because it could restore normal function. Conversely, an organism's two normal alleles would free this extra gene copy from having to maintain normal function. This would allow the gene to accumulate mutations without major consequence, which could ultimately lead to the formation of a new, unique gene. Copy number variations may contribute to evolution because they are yet another mechanism for organisms to achieve genetic innovation.

Connecting the Concepts

Basic research into the nature and organization of genes in various organisms allowed geneticists to produce recombinant DNA molecules. An understanding of transcription and translation made it possible for scientists to manipulate the expression of genes in organisms. These breakthroughs have ushered in a biotechnology revolution. Because the genetic code is almost universal, bacteria and eukaryotic cells are now used to produce vaccines, hormones, and growth factors for use in humans. Plants and animals are also engineered to make a product or to possess characteristics desired by humans. Biotechnology also offers the promise of treating and even some day curing human genetic disorders such as muscular dystrophy, cystic fibrosis, hemophilia, and many others. It also shows promise in creating hardier crops that could help alleviate food shortages in many parts of the world.

The growing field of genomics also shows promise. Now that the entire human genome has been sequenced, scientists can use this information to determine which genes function in particular cells and also to determine people's genetic profiles for the purpose of prescribing medications, diagnosing illness, and preventing future problems. Genetists have also sequenced the genomes of many other species, and comparisons between them is yielding valuable new insights into the relationships between species, impacting taxonomy and evolutionary biology.

summary

14.1 DNA Cloning

DNA cloning can isolate a gene and produce many copies of it. The gene can be studied in the laboratory or inserted into a bacterium, plant, or animal. Then, this gene may be transcribed and translated to produce a protein, which can become a commercial product or used as a medicine.

Two methods are currently available for making copies of DNA: recombinant DNA technology and the polymerase chain reaction (PCR). Recombinant DNA contains DNA from two different sources. A restriction enzyme is used to cleave plasmid DNA and to cleave foreign DNA. The resulting "sticky ends" facilitate the insertion of foreign DNA into vector DNA. The foreign gene is sealed into the vector DNA by DNA ligase. Both bacterial plasmids and viruses can be used as vectors to carry foreign genes into bacterial host cells.

PCR uses the enzyme DNA polymerase to quickly make multiple copies of a specific piece (target) of DNA. PCR is a chain reaction because the targeted DNA is replicated over and over again. Analysis of DNA segments following PCR has all sorts of uses from assisting genomic research to doing DNA fingerprinting for the purpose of identifying individuals and their paternity.

14.2 Biotechnology Products

Transgenic organisms have had a foreign gene inserted into them. Genetically modified bacteria, agricultural plants, and farm animals now produce commercial products of interest to humans, such as hormones and vaccines. Bacteria usually secrete the product. The seeds of plants and the milk of animals contain the product.

Transgenic bacteria have also been engineered to promote the health of plants, perform bioremediation, extract minerals, and produce chemicals. Transgenic crops, engineered to resist herbicides and pests, are commercially available. Transgenic animals have been given various genes, in particular the one for bovine growth hormone (bGH). Cloning of animals is now possible.

14.3 Gene Therapy

Gene therapy, by either ex vivo or in vivo methods, is used to correct the genotype of humans and to cure various human ills. Ex vivo gene therapy has apparently helped children with SCID lead normal lives. In vivo treatment for cystic fibrosis has been less successful. A number of in vivo therapies are being employed in the war against cancer and other human illnesses, such as cardiovascular disease.

14.4 Genomics

Researchers now know the sequence of all the base pairs along the length of the human chromosomes. So far, researchers have found only 20,500 human genes that code for proteins; the rest of our DNA consists of regions that do not code for a protein. Currently, researchers are placing an emphasis on functional and comparative genomics.

Genes only comprise 1.5% of the human genome. The rest of this DNA is surprisingly more active than once thought. About half of this DNA consists of repetitive DNA elements, which may be in tandem or interspersed throughout several chromosomes. Some of this DNA is made up of mobile DNA sequences called transposons, which are a driving evolutionary force within the genome. The remaining half of the genome remains unclassified, but even these unknown DNA sequences may play an important role in regulation of gene expression, and challenging the classical definition of the gene. Functional genomics aims to understand the function of protein coding regions and noncoding regions of our genome. To that end, researchers are utilizing new tools such as DNA microarrays. Microarrays can also be used to create an individual's genetic profile, which can be helpful in predicting illnesses and how a person will react to particular medications.

Comparative genomics has revealed that there is little difference between the DNA sequence of our bases and those of many other organisms. Genome comparisons have revolutionized our understanding of evolutionary relations by revealing previously unknown relationships between organisms.

Proteomics is the study of which genes are active in producing proteins in which cells under which circumstances. Bioinformatics is the use of the computer to assist proteomics and functional and comparative genomics.

understanding the terms

bioinformatics 259
biotechnology products 252
cloning 250
comparative genomics 258
complementary DNA (cDNA) 251
DNA ligase 250
DNA microarray 258
ex vivo gene therapy 254
functional genomics 258
gene cloning 250
gene pharming 253
gene therapy 250
genetically modified organism (GMO) 252
genetic profile 258
genomics 255
homologous gene 260

Human Genome Project (HGP) 255
intergenic sequence 256
interspersed repeat 257
in vivo gene therapy 254
plasmid 250
polymerase chain reaction (PCR) 250
proteome 258
proteomics 258
recombinant DNA (rDNA) 250
repetitive DNA element 256
restriction enzyme 250
short tandem repeat (STR) profiling 251
structural genomics 255
tandem repeat 256
transgenic organism 250
transposon 257
vector 250
xenotransplantation 254

Match the terms to these definitions:

a. _____ Bacterial agent that stops viral reproduction by cleaving viral DNA; used to cut DNA at specific points during production of recombinant DNA.
b. _____ Free-living organism in the environment that has had a foreign gene inserted into it.
c. _____ Use of animal organs, instead of human organs, in human transplant patients.
d. _____ Production of identical copies; in genetic engineering, the production of many identical copies of a gene.
e. _____ Biotechnology method that can quickly produce many duplicate copies of a piece of DNA.

reviewing this chapter

1. What is the methodology for producing recombinant DNA so useful for gene cloning? 250
2. What is the polymerase chain reaction (PCR), and how is it carried out to produce multiple copies of a DNA segment? 251–52
3. How does STR profiling produce a DNA fingerprint? 251–52
4. What are some practical applications of DNA segment analysis following PCR? 252
5. For what purposes have bacteria, plants, and animals been genetically altered? 252–54
6. Explain and give examples of ex vivo and in vivo gene therapies in humans. 254–55
7. What was the purpose of the Human Genome Project? What is the goal of functional genomics? 255–58
8. What insights into evolutionary relationships between organisms are arising from comparative genomics? 258
9. Describe the various types of intergenic DNA sequences found within the genome. 256–57
10. What are the goals of proteomics and bioinformatics? 258–60

testing yourself

Choose the best answer for each question.

1. Using this key, put the phrases in the correct order to form a plasmid-carrying recombinant DNA.

KEY:

(1) use restriction enzymes
(2) use DNA ligase
(3) remove plasmid from parent bacterium
(4) introduce plasmid into new host bacterium

a. 1, 2, 3, 4
b. 4, 3, 2, 1
c. 3, 1, 2, 4
d. 2, 3, 1, 4

2. Restriction enzymes found in bacterial cells are ordinarily used
   a. during DNA replication.
   b. to degrade the bacterial cell's DNA.
   c. to degrade viral DNA that enters the cell.
   d. to attach pieces of DNA together.

3. A genetic profile can
   a. assist in maintaining good health.
   b. be accomplished utilizing bioinformatics.
   c. show how many genes are normal.
   d. be accomplished utilizing a microarray.
   e. Both a and d are correct.

4. Bacteria are able to successfully transcribe and translate human genes because
   a. both bacteria and humans contain plasmid vectors.
   b. bacteria can replicate their DNA, but humans cannot.
   c. human and bacterial ribosomes are vastly different.
   d. the genetic code is nearly universal.

5. Bioinformatics can
   a. assist genomics and proteomics.
   b. compare our genome to that of a monkey.
   c. depend on computer technology.
   d. match up genes with proteins.
   e. All of these are correct.

6. The polymerase chain reaction
   a. uses RNA polymerase.
   b. takes place in huge bioreactors.
   c. uses a temperature-insensitive enzyme.
   d. makes lots of nonidentical copies of DNA.
   e. All of these are correct.

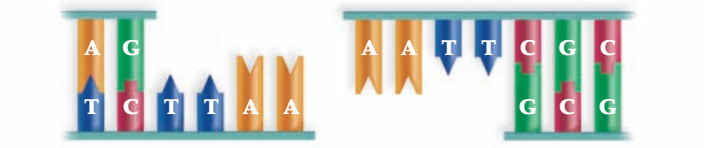
7. Which is a true statement?
   a. Genomics would be slow going without bioinformatics.
   b. Genomics is related to the field of proteomics.
   c. Genomics has now moved on to functional and comparative genomics.
   d. Genomics shows that we are related to all other organisms tested so far.
   e. All of these are correct.

8. DNA amplified by PCR and then used for fingerprinting could come from
   a. any diploid or haploid cell.
   b. only white blood cells that have been karyotyped.
   c. only skin cells after they are dead.
   d. only purified animal cells.
   e. Both b and d are correct.

9. Which was used to find the function of the cystic fibrosis gene?
   a. microarray
   b. proteomics
   c. comparative genomics and bioinformatics
   d. sequencing the gene

10. Which of these pairs is incorrectly matched?
    a. DNA ligase—mapping human chromosomes
    b. protoplast—plant cell engineering
    c. DNA fragments—DNA fingerprinting
    d. DNA polymerase—PCR

11. Which matches best to proteomics?
    a. Start with known gene sequences and build proteins.
    b. Use a microarray to discover what proteins are active in particular cells.
    c. Use bioinformatics to discover the proteins in the cells of other organisms.
    d. Match up known proteins with known genes.
12. Which is not a correct association with regard to bioengineering?
    a. plasmid as a vector—bacteria
    b. protoplast as a vector—plants
    c. RNA virus as a vector—human stem cells
    d. All of these are correct.
13. Proteomics is used to discover
    a. what genes are active in what cells.
    b. what proteins are active in what cells.
    c. the structure and function of proteins.
    d. how proteins interact.
    e. All but a are correct.
14. Which of these is an incorrect statement?
    a. Bacteria usually secrete the biotechnology product into the medium.
    b. Plants are being engineered to have human proteins in their seeds.
    c. Animals are engineered to have a human protein in their milk.
    d. Animals can be cloned, but plants and bacteria cannot.
15. Repetitive DNA elements
    a. may be tandem or spread across several chromosomes.
    b. are found in centromeres and telomeres.
    c. make up nearly half of human chromosomes.
    d. may be present just a few to many thousands of copies.
    e. All of these are correct.
16. Because of the Human Genome Project, we now know
    a. the sequence of the base pairs of our DNA.
    b. the sequence of all genes along the human chromosomes.
    c. all the mutations that lead to genetic disorders.
    d. All of these are correct.
    e. Only a and c are correct.
17. Which of the following delivery methods is not used in gene therapy?
    a. virus
    b. nasal sprays
    c. liposomes
    d. electric currents
18. The restriction enzyme called *Eco*RI has cut double-stranded DNA in the following manner. The piece of foreign DNA to be inserted has what bases from the left and from the right?



19. Which of these is a true statement?
    a. Plasmids can serve as vectors.
    b. Plasmids can carry recombinant DNA, but viruses cannot.
    c. Vectors carry only the foreign gene into the host cell.
    d. Only gene therapy uses vectors.
    e. Both a and d are correct.
20. Gene therapy
    a. is sometimes used in medicine today.
    b. is always successful.
    c. is only used to cure genetic disorders, such as SCID and cystic fibrosis.
    d. makes use of viruses to carry foreign genes into human cells.
    e. Both a and d are correct.

thinking scientifically

1. Before the human genome was sequenced, gene discovery was accomplished through the use of DNA libraries. A genomic library is a set of cloned DNA segments that altogether are representative of the genome of an organism, whereas a cDNA library contains only expressed DNA sequences for a particular cell. How might these libraries be used to map the introns and exons of a gene within the genome?
2. The Science Focus on page 260 describes copy number variations within the genome. Copy number variations do not always contain genes. How might having extra or missing copies of intergenic DNA sequences be beneficial? How might it be harmful?

bioethical issue

Transgenic Crops

Transgenic plants can possibly allow crop yields to keep up with the ever-increasing worldwide demand for food. And some of these plants have the added benefit of requiring less fertilizer and/or pesticides, which are harmful to human health and the environment.

But some scientists believe transgenic crops pose their own threat to the environment, and many activists believe transgenic plants are themselves dangerous to our health. Studies have shown that wind-carried pollen can cause transgenic crops to hybridize with nearby weedy relatives. Although it has not happened yet, some fear that characteristics acquired in this way might cause weeds to become uncontrollable pests. Or perhaps a toxin produced by transgenic crops could possibly hurt other organisms in the field. Many researchers are conducting tests to see if this might occur. And although transgenic crops have not caused any illnesses in humans so far, some scientists concede the possibility that people could be allergic to the transgene's protein product. After unapproved genetically modified corn was detected in Taco Bell taco shells several years ago, a massive recall pulled about 2.8 million boxes of the product from grocery stores.

Already, transgenic plants must be approved by the Food and Drug Administration before they are considered safe for human consumption, and they must meet certain Environmental Protection Administration standards. Some people believe safety standards for transgenic crops should be further strengthened, while others fear stricter standards will result in less food produced. Another possibility is to retain the current standards but require all biotech foods to be clearly labeled so the buyer can choose whether or not to eat them. Which approach do you prefer?

Biology website

The companion website for *Biology* provides a wealth of information organized and integrated by chapter. You will find practice tests, animations, videos, and much more that will complement your learning and understanding of general biology.

http://www.mhhe.com/maderbiology10

part III

Evolution

Evolution may seem like a foreign topic to you, but it need not be. Evolution simply refers to the changes that occur as one generation begets the next generation. Just as you can trace your ancestry to your great-grandparents and beyond, so all of life can trace its ancestry to the first living cell or cells. A remarkable finding in recent times has been that some of our genes are even the same as those of prokaryotes. This couldn't be if we were not related to prokaryotes.

The changes that occur as a population reproduces assist its members in finding food, mates, a place to live, and even in avoiding dangers. Consider, for example, that resistant bacteria are able to avoid the danger of being killed by an antibiotic. When an antibiotic is taken for a staph infection, a few staph bacteria may be able to survive, and they are the ones that will produce the next generation of staph bacteria. Soon, most members of future generations are resistant. This is the manner in which all adaptations to the environment occur. From carnivorous plants to bats and whales, all life is adapted to its particular environment.

Evolution, the topic of this part, explains the world of living things; how it came to be, and why it is so diverse. The next time you are in a natural area, observe the diversity and say, "Evolution!"

15

Darwin and Evolution

In 2006, a fossil snake was discovered with a pelvic girdle and legs. Charles Darwin, sometimes called the father of evolution, would have been pleased because such a fossil shows us how evolution occurred. Living pythons have leg remnants, so now we have both living and fossil evidence that snakes had legged ancestors.

Darwin's most noted contribution was to discover what caused the tree of life to have so many branches. Much data allowed him to conclude that nature (the environment) selects which members of a population will reproduce and pass on their adaptive traits to their offspring. Life is diverse because environments are diverse. This chapter is about Darwin's contribution to the field of evolution. First, we take a look at evolutionary thought when Darwin began his work. Then, we retrace Darwin's trip around the world and present still other data that allowed Darwin to develop his theory of evolution by natural selection. Finally, we take a look at evidence that supports Darwin's statement that evolution is "common descent with modification."

Today's green tree python (*Morelia viridis*) has vestigial legs. This tells us that an ancestor of snakes most likely had functional legs.

concepts

15.1 History of Evolutionary Thought

In December 1831, a new chapter in the history of biology had its humble origins. A 22-year-old naturalist, Charles Darwin (1809–82), set sail on a journey of a lifetime aboard the British naval vessel the HMS *Beagle* (Fig. 15.1). Darwin's primary mission on his journey around the world was to expand the navy's knowledge of potential natural resources, such as water and food in foreign lands. Prior to Darwin, the worldview was forged by deep-seated beliefs that were held to be intractable truths and not by experimentation and observation of the natural word. In contrast, Darwin used a variety of data to come to the conclusion that the Earth is very old, not young, and that biological evolution is the method by which species arise and change. The acceptance of the Darwinian view of the world was fostered by a scientific and intellectual revolution that occurred in both the scientific and social realms of the mid-1800s.

Although it is often believed that Darwin (Fig. 15.2) forged this change in worldview by himself, several biologists during the preceding century and some of Darwin's con-

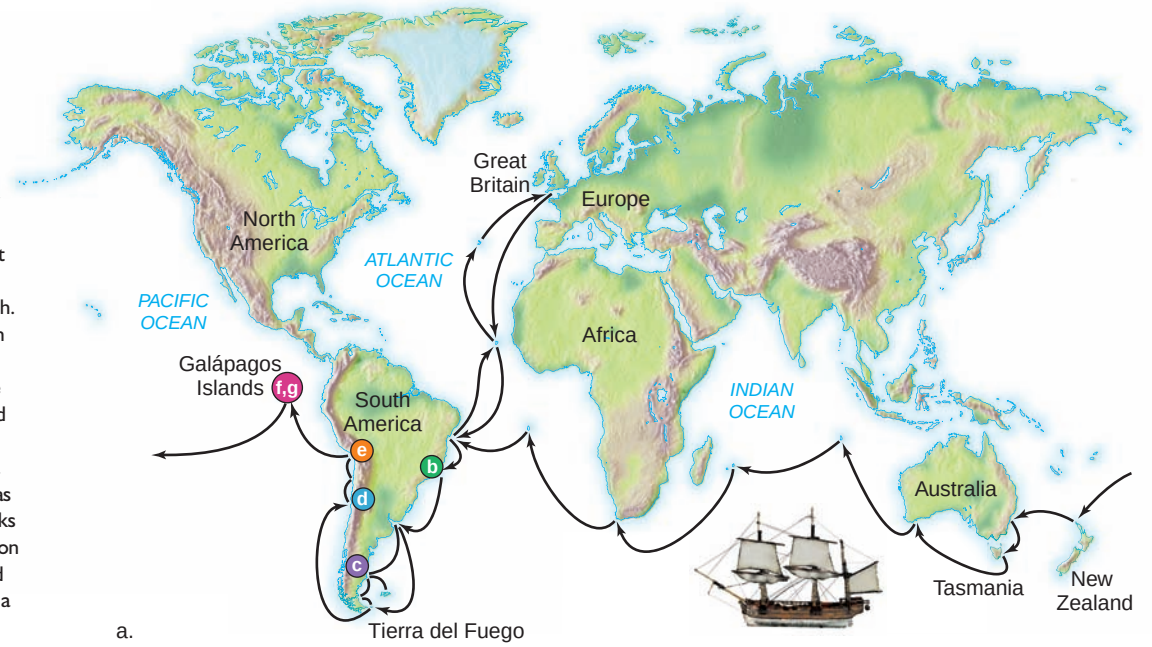


FIGURE 15.1

Voyage of the HMS *Beagle*.

a. Map shows the journey of the HMS *Beagle* around the world. Notice that the map's encircled colors are keyed to the encircled colors in the photographs, which show us what Charles Darwin may have observed in or near South America. **b.** As Darwin traveled along the east coast of South America, he noted that a bird called a rhea looked like the African ostrich. **c.** The sparse vegetation of the Patagonian Desert is in the southern part of the continent. **d.** The Andes Mountains of the west coast have strata containing fossilized organisms. **e.** The lush tropical rain forest contains a high and unique diversity of life. **f.** On the Galápagos Islands, marine iguanas have large claws to help them cling to rocks and blunt snouts for eating algae growing on rocks. **g.** Galápagos finches are specialized to feed in various ways. This finch is using a cactus spine to probe for insects.

temporaries slowly began to accept the idea that species change over time. This concept would eventually be known as **evolution** [L. *evolutio*, an unrolling] and is now considered the unifying principle of all the biological sciences. Evolution explains both the unity and diversity of life on Earth. First, evolution illustrates that living things share like characteristics because they have a common ancestry. Evolution also explains how species adapt to various habitats and ways of life with the result that life is very diverse.

The history of evolutionary thought is a history of ideas about descent and adaptation. Darwin used the expression "descent with modification," by which he meant that as descent through generations occurs over time, so does diversification. Darwin brilliantly saw the process of adaptation as a means by which the diversity of species arises.

FIGURE 15.2 Charles Darwin (1809–82).

The theory of evolution is usually identified with Charles Darwin, who, along with Alfred Wallace, proposed a mechanism for evolution. This portrait of Darwin dates from 1831; he did not publish his authoritative book, *On the Origin of Species*, until 1859.

Mid-Eighteenth-Century Contributions

Taxonomy, the science of classifying organisms, was an important endeavor during the mid-eighteenth century. Chief among the taxonomists was Carolus Linnaeus (1707–78), who developed the binomial system of nomenclature (a two-part name for species, such as *Homo sapiens*) and a system of classification for living things. In addition to taxonomy, comparative anatomy, the evaluation of similar structures across a variety of species, was of interest to biologists prior to Darwin. By the late eighteenth century, scientists had discovered fossils and knew that they were the remains of plants and animals from the past. Explorers traveled the world bringing back not only **extant** (living species) but also fossils to be compared to living species. At first, they believed that each type fossil had a living descendant, but eventually some fossils did not seem to match up well with known species. Baron Georges Cuvier was the first to suggest that some species known only from the fossil record had become extinct.

Linnaeus, like other taxonomists of his time, believed in the fixity of species. Each species had an "ideal" structure and function and also a place in the *scala naturae*, a sequential ladder of life. The simplest and most material being was on the lowest rung of the ladder, and the most complex and spiritual being was on the highest rung. In this view, human beings occupied the highest rung of the ladder. These ideas, which were consistent with Judeo-Christian teachings about special creation, can be traced to the works of the famous Greek philosophers Plato (427–347 BC) and Aristotle (384–322 BC). Plato said that every object on Earth was an imperfect copy of an ideal form, which can be deduced upon reflection and study. To Plato, individual variations were imperfections that only distract the observer. Aristotle saw that organisms were diverse and some were more complex than others. His belief that all organisms could be arranged in order of increasing complexity became the *scala naturae* just described.

Linnaeus and other taxonomists wanted to determine the ideal characteristics of each species and also wanted to discover the proper rank for each species in the *scala naturae*. Therefore, for most of his working life, Linnaeus did not even consider the possibility of evolutionary change. There is evidence, however, that he did eventually perform hybridization experiments, which made him think that a species might change with time.

Georges-Louis Leclerc, better known by his title, Count Buffon (1707–88), was a French naturalist who devoted many years of his life to writing a 44-volume natural history series of texts that described all known plants and animals. He provided evidence of descent with modification, and he even speculated on various causative mechanisms such as environmental influences, migration, geographic isolation, and the struggle for existence. Buffon seemed to waver, however, as to whether or not he recognized evolutionary descent, and often he professed to believe in special creation and the fixity of species.

Erasmus Darwin (1731–1802), Charles Darwin's grandfather, was a physician and a naturalist. His writings on both botany and zoology, although they were mostly in footnotes and asides, contained many comments that suggested the possibility of common descent. He based his conclusions on changes undergone by animals during development, artificial selection by humans, and the presence of **vestigial structures** (structures or organs that are thought to have been functional in an ancestor but are reduced and nonfunctional in a descendant). Like Buffon, Erasmus Darwin offered no conclusive mechanism by which evolutionary descent might occur.

Late-Eighteenth-/Early-Nineteenth-Century Contributions

Even though Linnaeus was never an evolutionist, his hierarchical method of classifying organisms is consistent with modern evolutionary thinking. This is the reason that Linnaeus' basic method of classification has been modified even until today. Whether Linnaean classification will always be flexible enough for continual modification is a question that is even now being asked.

Cuvier and Catastrophism

Baron Georges Cuvier (1769–1832), a distinguished zoologist, used comparative anatomy to develop a system of classifying animals. He also founded the science of

paleontology [Gk. *palaios,* old; *ontos,* having existed; *-logy,* study of], the study of fossils, and was quite skilled at using fossil bones to deduce the structure of an animal (Fig. 15.3*a*).

Because Cuvier was a staunch advocate of special creation and the fixity of species, he faced a real problem when a particular region showed a succession of life-forms in the Earth's strata (layers). To explain these observations, he hypothesized that a series of local catastrophes or mass extinctions had occurred whenever a new stratum of that region showed a new mix of fossils. After each catastrophe, the region was repopulated by species from surrounding areas, and this accounted for the appearance of new fossils in the new stratum. The result of all these catastrophes was change appearing over time. Some of Cuvier's followers even suggested that there had been worldwide catastrophes and that after each of these events, God created new sets of species. This explanation of the history of life came to be known as **catastrophism** [Gk. *katastrophe,* calamity, misfortune].

Lamarck and Acquired Characteristics

Jean-Baptiste de Lamarck (1744–1829) was the first biologist to offer a mechanism for how evolution occurs and to link diversity with adaptation to the environment. Lamarck's ideas about descent were entirely different from those of Cuvier, perhaps because Lamarck specialized in the study of invertebrates (animals without backbones), while Cuvier was a vertebrate zoologist, who studied animals with backbones. Lamarck concluded, after studying the succession of life-forms in strata, that more complex organisms are descended from less complex organisms. He mistakenly said, however, that increasing complexity is the result of a natural force—a desire for perfection—that is inherent in all living things.

To explain the process of adaptation to the environment, Lamarck supported the idea of **inheritance of acquired characteristics**—that the environment can bring about inherited change. One example that he gave—and for which he is most famous—is that the long neck of a giraffe developed over time because animals stretched their necks to reach food in tall trees and then passed on a long neck to their offspring (Fig. 15.3*b*). His hypothesis of the inheritance of acquired characteristics has never been substantiated by experimentation. The molecular mechanism of inheritance explains why. Phenotypic changes acquired during an organism's lifetime do not result in genetic changes that can be passed to subsequent generations.

Check Your Progress 15.1

1. According to Lamarck, if someone dyed their hair dark to make their blonde less visible at night, what color hair would be passed on to their offspring?
2. Based on Cuvier's ideas, if an asteroid were to impact the Earth and cause 99% of living species to go extinct, where would replacement species come from?

a.

b.

FIGURE 15.3 Evolutionary thought before Darwin.

a. Cuvier reconstructed animals such as extinct mastodons and said that catastrophes followed by repopulations could explain why species change over time. **b.** Lamarck explained the long neck of a giraffe according to his ideas about the inheritance of acquired characteristics.

15.2 Darwin's Theory of Evolution

When Darwin signed on as the naturalist aboard the HMS *Beagle*, he possessed a suitable background for the position. Since childhood, he was a devoted student of nature and a collector of insects. At age 16, Darwin was sent to medical school to follow in the footsteps of his grandfather and father. However, his sensitive nature prevented him from studying medicine, and he enrolled in the school of divinity at Christ College at Cambridge, intent on becoming a clergyman.

While at Christ College, he attended many lectures in biology and geology to satisfy his interest in natural science. During this time, he became the protégé of his teacher and friend, Reverend John Henslow. Darwin gained valuable experience in geology in the summer of 1831, conducting fieldwork with Adam Sedgewick. Shortly after Darwin was awarded his BA, Henslow recommended him to serve, without pay, as ship's naturalist on the HMS *Beagle*. The trip was to take two years—but ended up taking five years—and the ship was to traverse the Southern Hemisphere (see Fig. 15.1), where life is most abundant and varied. Along the way, Darwin encountered forms of life very different from his native England.

Occurrence of Descent

Although it was not his original intent, Darwin began to gather evidence that organisms are related through common descent and that adaptation to various environments results in diversity. Darwin also began contemplating the "mystery of mysteries," the origin of new species.

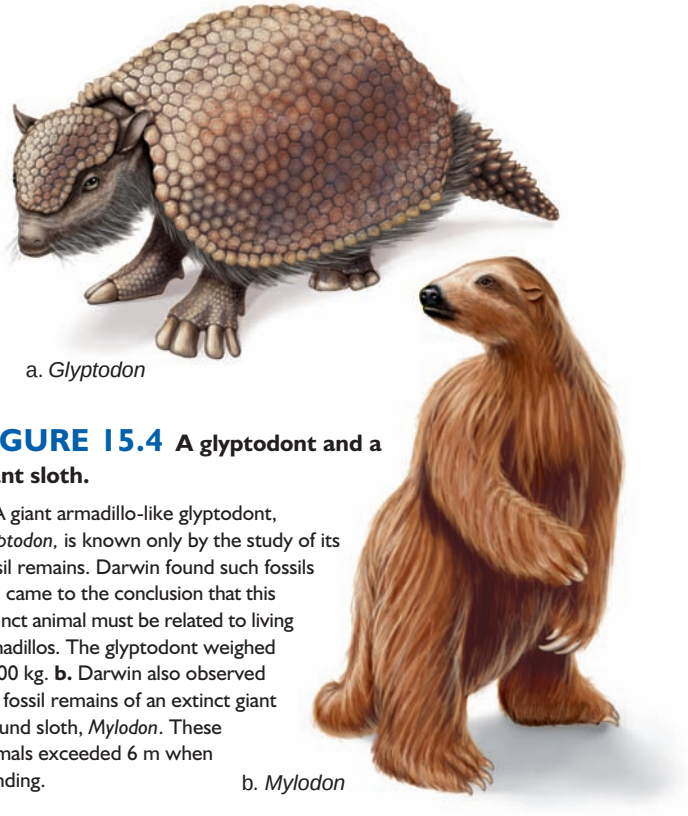
Geology and Fossils

Darwin took Charles Lyell's *Principles of Geology* on the voyage, which presented arguments to support a theory of geological change proposed by James Hutton. In contrast to the catastrophists, Hutton explained that the Earth was subject to slow but continuous cycles of rock formation through erosion and uplift. Weather causes erosion; thereafter, dirt and rock debris are washed into the rivers and transported to the oceans. These loose sediments are deposited in thick layers, which are converted eventually into sedimentary rocks. These sedimentary rocks, which often contain fossils, are then uplifted from below sea level to form land. Hutton concluded that extreme geological changes can be accounted for by slow, natural processes, given enough time. Lyell went on to propose the theory of **uniformitarianism,** which stated that these slow changes occurred at a uniform rate and that the natural processes witnessed today are the same processes that occurred in the past. Hutton's general ideas about slow and continual geological change are still accepted today, although modern geologists realize that rates of change have not always been uniform through history. Darwin was not taken by the idea of uniform change, but he was convinced, as was Lyell, that the Earth's massive geological changes are the result of extremely slow processes and that the Earth, therefore, must be very old.

On his trip, Darwin observed massive geological changes firsthand. When he explored what is now Argentina, he saw raised beaches for great distances along the coast. In Chile, he witnessed the effects of an earthquake that caused the land to rise several feet, leaving marine shells inland, well above sea level. This observation, along with marine shells high in the cliffs of the impressive Andes Mountains, supported Lyell's theory of slow geologic changes of a very old planet. While Darwin was making geological observations, he also collected fossil specimens. For example, on the east coast of South America, he found the fossil remains of an armadillo-like animal (*Glyptodon*), the size of a small modern-day car, and giant ground sloths, the smallest of which stood nearly 3 m tall (Fig. 15.4). Once Darwin accepted the supposition that the Earth must be very old, he began to think that there would have been enough time for descent with modification to occur. Therefore, living forms could be descended from extinct forms known only from the fossil record. It would seem that species were not fixed; instead, they changed over time.

Biogeography

Biogeography [Gk. *bios,* life, *geo,* earth, and *grapho,* writing] is the study of the range and geographic distribution of lifeforms on Earth. Darwin could not help but compare the animals of South America to those with which he was familiar. For example, instead of rabbits, he found the Patagonian hare in the grasslands of South America. The Patagonian hare has long legs and ears but the face of a guinea pig, a



a. *Glyptodon*

b. *Mylodon*

FIGURE 15.4 A glyptodont and a giant sloth.

a. A giant armadillo-like glyptodont, *Glyptodon,* is known only by the study of its fossil remains. Darwin found such fossils and came to the conclusion that this extinct animal must be related to living armadillos. The glyptodont weighed 2,000 kg. **b.** Darwin also observed the fossil remains of an extinct giant ground sloth, *Mylodon*. These animals exceeded 6 m when standing.

FIGURE 15.5

The European hare, (head only), and the Patagonian hare.

Lepus europaeus

Dolichotis patagonium

rodent also native to South America (Fig. 15.5). Did the Patagonian hare resemble a rabbit because the two types of animals were adapted to the same type of environment? Both animals ate grass, hid in bushes, and moved rapidly using long hind legs. Did the Patagonian hare have the face of a guinea pig because of common descent with guinea pigs?

As he sailed southward along the eastern coast of South America, Darwin saw how similar species replaced one another. For example, the greater rhea (an ostrichlike bird) found in the north was replaced by the lesser rhea in the south. Therefore, Darwin reasoned that related species could be modified according to environmental differences (i.e., Northern vs. Southern latitudes). When he explored the Galápagos Islands, he found further evidence of this phenomenon. The Galápagos Islands are a small group of volcanic islands formed 965 km off the western coast of South America when underwater volcanoes emerged from the ocean. These islands are too far from the mainland for most terrestrial animals and plants to colonize, yet life was present there. The types of plants and animals found there were slightly different from species Darwin had observed on the mainland and, even more important, they also varied from island to island. Where did animals and plants inhabiting these islands come from, and why were different species on each island?

Tortoises. Each of the Galápagos Islands seemed to have its own type of tortoise, and Darwin began to wonder if this could be correlated with a difference in vegetation among the islands (Fig. 15.6). Long-necked tortoises seemed to inhabit only dry areas, where food was scarce, most likely because the longer neck was helpful in reaching tall-growing cacti. In moist regions with relatively abundant ground foliage, short-necked tortoises were found. Had an ancestral tortoise from the mainland of South America given rise to these different types, each adapted to a different environment?

Finches. Darwin almost overlooked the finches because of their unassuming nature compared with many of the other animals in the Galápagos. However, these birds would eventually play a major role in his thoughts about geographic isolation. The finches of the Galápagos Islands seemed to Darwin like mainland finches, but they exhibited significant variety with regard to their beaks (see Fig. 15.10). Today, there are ground-dwelling finches with beaks sized appropriate to the seeds they feed on, tree-dwelling finches with beaks sized according to their insect prey, and a cactus-eating finch with a more pointed beak. The most unusual of the finches is a woodpecker-type finch. This bird has a sharp beak to chisel through tree bark but lacks the long tongue characteristic of a true woodpecker, which probes for insects. To compensate for this, the bird carries a twig or cactus spine in its beak and uses it to poke into crevices (see Fig. 15.1*g*). Once an insect emerges, the finch drops this tool and seizes the insect with its beak.

Later, Darwin speculated as to whether these different species of finches could have descended from a type of mainland finch. In other words, he wondered if a finch from South America was the common ancestor to all the types on the Galápagos Islands. Had speciation occurred because the islands allowed isolated populations of birds to evolve independently? Could the present-day species have resulted from accumulated changes occurring within each of these isolated populations?

a.

b.

FIGURE 15.6 Galápagos tortoises.

Darwin wondered if all of the Galápagos tortoises, *Geochelone nigra,* of the various islands were descended from a common ancestor. **a.** The tortoises with dome shells and short necks feed at ground level and occur on well-watered islands where grass is available. **b.** Those with shells that flare up in front have long necks and are able to feed on tall, treelike cacti. They live on arid islands where prickly pear cacti are the main food source. Only on these islands are the cacti treelike.

Natural Selection and Adaptation

Upon returning to England, Darwin began to reflect on the voyage of the HMS *Beagle* and to collect evidence supporting his ideas about how organisms adapt to the environment. Darwin decided that adaptations develop over time (instead of being the instant work of a creator), and he began to think about a mechanism that would allow this to happen. In the late 1850s, Darwin proposed **natural selection** as a mechanism for evolutionary change. Meanwhile, Alfred Russel Wallace, who is discussed in the Science Focus on page 274, proposed the same concept based on similar observations from the other side of the globe. Natural selection is a process consisting of these components:

- The members of a population have inheritable variations. Variation within a population of a species occurs for a multitude of traits, many of which are inheritable.
- A population produces more offspring in each generation than the environment can support. In any given environment, there is a limited amount of food, water, physical space, and other resources for which individuals must compete.
- Some individuals have favorable traits that enable them to better compete for limited resources. The individuals with favorable traits acquire more resources than the individuals with less favorable traits and can devote more energy to reproduction. Darwin called this ability to have more offspring *differential reproductive success.*
- Natural selection can result in a population adapted to the local environment. An increasing proportion of individuals in each succeeding generation will have the favorable characteristics—characteristics suited to surviving and reproducing in that environment. In this way, evolution brings about adaptation to the environment.

FIGURE 15.7 Variation in a population.

For Darwin, variations, such as those seen in a human population, are highly significant and are required for natural selection to result in adaptation to the environment.

Each of these steps leading to adaptation is now examined in more detail.

Organisms Have Inheritable Variations

Darwin emphasized that the members of a population vary in their functional, physical, and behavioral characteristics (Fig. 15.7). Before Darwin, variations were viewed as imperfections that should be ignored since they were not important to the description of a species. Darwin emphasized that variations were essential to the natural selection process. He suspected—but did not have the evidence we have today—that the occurrence of variations is completely random; they arise by accident and for no particular purpose. New variations are just as likely to be harmful as helpful to the organism.

The variations that make adaptation to the environment possible are those that are passed on from generation to generation. The science of genetics was not yet well established, so Darwin was never able to determine the cause of variations or how they are passed on. Today, we realize that genes, along with the environment, determine the phenotype of an organism, and that mutations and recombination of alleles during sexual reproduction can cause new variations to arise.

Natural selection can only operate on variations that are already available in the population's gene pool; it lacks any directedness towards "improvement" or anticipation of future environmental changes—and the environment of living things is constantly changing.

Organisms Compete for Resources

In Darwin's time, a socioeconomist, Thomas Malthus, stressed the reproductive potential of human beings. He proposed that death and famine were inevitable because the human population tends to increase faster than the supply of food. Darwin applied this concept to all organisms and saw that the available resources were not sufficient for all members of a population to survive. He calculated the reproductive potential of elephants, assuming an average life span of 100 years and a breeding span from 30–90 years. Given these assumptions, a single female probably bears no fewer than six young, and if all these young survive and continue to reproduce at the same rate, after only 750 years, the descendants of a single pair of elephants will number about 19 million! Obviously, no environment has the resources to support an elephant population of this magnitude, and no such elephant population has ever existed. This overproduction potential of a species is often referred to as Darwin's geometric ratio of increase.

Organisms Differ in Reproductive Success

Fitness is the reproductive success of an individual relative to other members of a population. The most-fit individuals are the ones that capture a disproportionate amount of resources, and convert these resources into a larger number of viable offspring. Since organisms vary anatomically and physiologically and the challenges of local environments vary, what determines fitness varies for different populations. For example, among western diamondback rattlesnakes (*Crotalus atrox*)

FIGURE 15.8 **Artificial selection of animals.**

All dogs, *Canis lupus familiaris,* are descended from the gray wolf, *Canis lupus,* which began to be domesticated about 14,000 years ago. The process of diversification has led to extreme phenotypic differences. Several factors may have contributed: (1) The wolves under domestication were separated from other wolves because human settlements were separate, and (2) humans in each settlement selected for whatever traits appealed to them. Artificial selection of dogs continues even today.

living on lava flows, the most fit are those that are black in color. But among those living on desert soil, the most fit are those with the typical light coloring with brown blotching. Background matching helps an animal both capture prey and avoid being captured; therefore, it is expected to lead to survival and increased reproduction.

In nature, interactions with the environment determine which members of a population reproduce to a greater degree than other members. In contrast to artificial selection, the result of natural selection is not predesired. Natural selection occurs because certain members of a population happen to have a variation that allows them to survive and reproduce to a greater extent than other members. For example, a variation that reduces water loss is beneficial to a desert plant; and one that increases the sense of smell helps a wild dog find its prey. Therefore, we expect organisms with these traits to have increased fitness.

Artificial Selection. Darwin noted that when humans help carry out **artificial selection,** the process by which a breeder chooses which traits to perpetuate, they select the animals and plants that will reproduce. For example, prehistoric humans probably noticed desirable variations among wolves and selected particular individuals for breeding. Therefore, the desired traits increased in frequency in the next generation. This same process was repeated many times over. The result today is the existence of many varieties of dogs, all descended from the wolf (Fig. 15.8). To take a modern example, foxes are very shy and normally shun the company of people, but in forty years time, Russian scientists have produced silver foxes that now allow themselves to be petted and even seek attention. They did this by selecting the most docile animals to reproduce. The scientists noted that some physical characteristics changed as well. The legs and tails became shorter, the ears become floppier, and the coat color patterns changed. Artificial selection is only possible because the original population exhibits a range of characteristics, allowing humans to select which traits they prefer to perpetuate. Therefore, several varieties of vegetables can be traced to a single ancestor. Chinese cabbage, brussel sprouts, and kohlrabi are all derived from a single species, *Brassica oleracea* (Fig. 15.9).

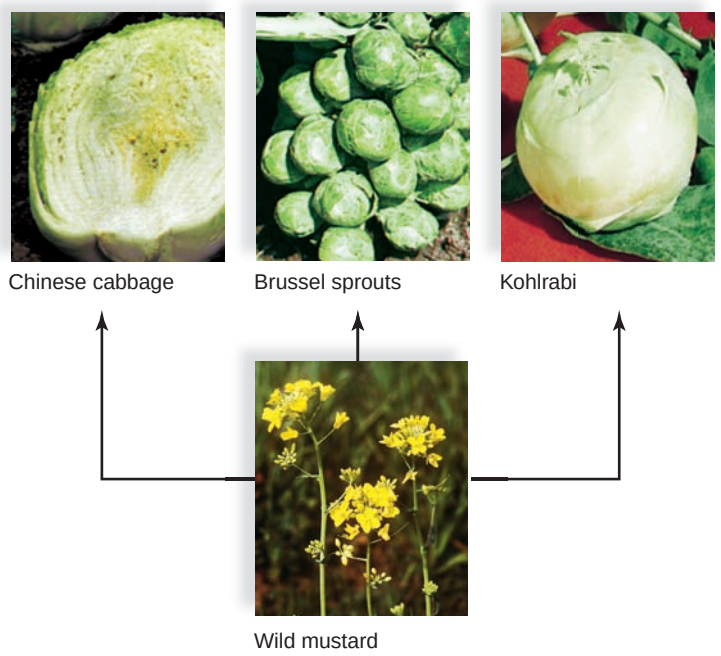


FIGURE 15.9 Artificial selection of plants.

The vegetables Chinese cabbage, brussel sprouts, and kohlrabi are derived from wild mustard, *Brassica oleracea.* Darwin described artificial selection as a model by which to understand natural selection. With natural selection, however, the environment and not human selection, provides the selective force.

Organisms Become Adapted

An **adaptation** [L. *ad,* toward, and *aptus,* fit, suitable] is a trait that helps an organism be more suited to its environment. Adaptations are especially recognizable when unrelated organisms, living in a particular environment, display similar characteristics. For example, manatees, penguins, and sea turtles all have flippers, which help them move through the water. In Chapter 1, we described other ways in which rockhopper penguins are adapted to their environment. Similarly, it can be noted that a Venus flytrap, a plant that lives in the nitrogen-poor soil of a bog, is able to catch and digest flies because it has leaves adapted to catching them.

Such adaptations to their specific environments result from natural selection. Differential reproduction generation after generation can cause adaptive traits to be increasingly represented in each succeeding generation. There are other processes of evolution aside from natural selection (see pages 286–93), but natural selection is the only process that results in adaptation to the environment.

On the Origin of Species by Darwin

After the HMS *Beagle* returned to England, Darwin waited more than 20 years to publish his ideas. During the intervening years, he used the scientific process to support his hypothesis that today's diverse life-forms are descended from a common ancestor and that natural selection is a mechanism by which species can change and new species can arise. In other words, Darwin hypothesized that new species could arise by gradual changes over time. Thus, when Darwin first published data supporting evolution by natural selection, he called the book *On the Origin of Species.*

Darwin was prompted to publish after receiving a letter from Alfred Russel Wallace in which Wallace also proposed natural selection as a mechanism for evolutionary change. The Science Focus on page 274 tells of this and the many other accomplishments of Wallace. Darwin was stunned to see his own theory being proposed by another and immediately went into print. Many feel that Wallace has not received the credit he deserves and that he should be given equal billing to Darwin for also discovering the mechanism by which evolution occurs. However, Darwin worked for many years to gather data to support his theory of natural selection. He later expanded *On the Origin of Species* and published over fifteen additional treatises with examples over the next two decades. Natural selection is so well supported that today we speak of the theory of evolution by natural selection.

science focus

Alfred Russel Wallace

Alfred Russel Wallace (1823–1913) is best known as the English naturalist who independently and simultaneously proposed natural selection as a mechanism for evolution (Fig. 15A). While working as a schoolteacher in his early twenties, he met the entomologist Henry Walter Bates, who interested him in insects. Together, they took a collecting trip to the Amazon, which lasted for several years. Wallace's knowledge of the world's diversity was further expanded by a tour of the Malay Archipelago from 1854–62. After studying the animals of every important island in the region, he divided the islands into a western group, with organisms like those found in Asia, and an eastern group, with organisms like those of Australia. The sharp line dividing these two island groups within the archipelago is now known as Wallace's Line (Fig. 15B). A narrow, but deep, strait occurs at Wallace's Line. At times during the past 50 million years, sea levels have lowered and land bridges have appeared between some of the islands. The strait, however, has always remained as a way to prevent animal dispersal between the western and eastern groups of islands.

In 1855, Wallace wrote an essay entitled "On the Law Which Has Regulated the Introduction of New Species," which stated "every species has come into existence coincident both in time and space with a preexisting closely allied species." It is clear, then, that by this date Wallace saw that species share a common ancestry and change over time. It was not until 1858, while suffering an attack of malaria, that he concluded changes in species are due to changes in the environment through natural selection. These conclusions were written in an essay to Darwin for comment. Darwin was stunned upon its receipt. Here before him was the hypothesis he had formulated as early as 1844. For 14 years, he collected copious data supporting natural selection as a mechanism for evolutionary change. Although a draft of a book was in hand by 1856, Darwin had never dared to publish it because he feared the criticism it would most likely receive. He told his friend and colleague Charles Lyell that Wallace's ideas were so similar to his own that even Wallace's "terms now stand as heads of my chapters."

Darwin suggested that Wallace's paper be published immediately, even though Darwin himself as yet had nothing in print. However, Lyell and others who knew of Darwin's detailed work substantiating the process of natural selection suggested that a joint paper be read to the Linnean Society. The title of Wallace's section was "On the Tendency of Varieties to Depart Indefinitely from the Original Type." Darwin presented an abstract of a paper he had written in 1844 and an abstract of his book *On the Origin of Species*. On July 1, 1858, two authors announced to the world that species evolve via natural selection. One year later, Darwin overshadowed Wallace, who was still in the field at the time, by publishing his famous book. However, many still referred to Wallace as "England's Greatest Living Naturalist" because of his diverse contributions well past the age of 90.

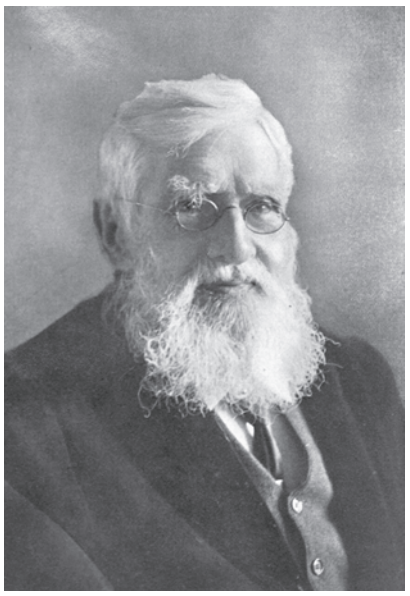


FIGURE 15A Alfred Russel Wallace.

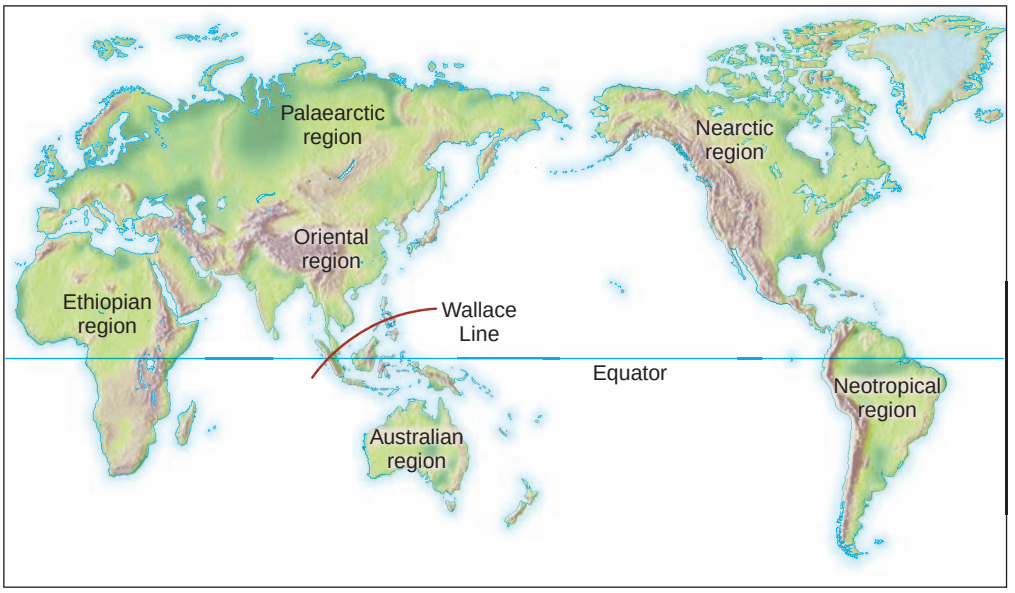


FIGURE 15B Wallace's Line.

Aside from presenting a hypothesis that natural selection explains the origin of new species, Wallace is well known for recognizing the sharp change in animal species inhabiting the islands on either side of a narrow strait bisecting the Malay Archipelago. This deep channel between the Oriental and Australian regions is called Wallace's Line, and serves as an impassable barrier to animal dispersal. By linking geography (the study of maps) to zoology (the study of animals), Wallace is often considered the "Father of Zoogeography."