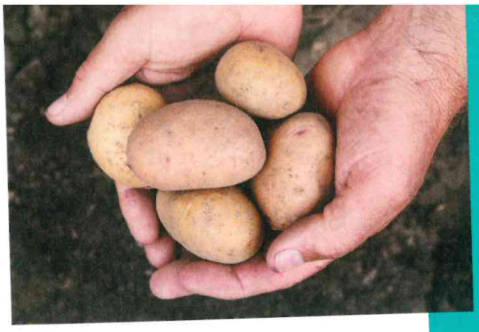
A Technology



ARE MORE GENES BETTER? YOU HAVE ABOUT THE SAME NUMBER OF GENES AS A MICROSCOPIC WORM, AND ONLY HALF AS MANY AS A RICE PLANT.

SOMEDAY SOON, GENETICALLY MODIFIED POTATOES COULD SAVE TENS OF THOUSANDS OF CHILDREN FROM DEATH BY CHOLERA.

latters

BIOLO

Using

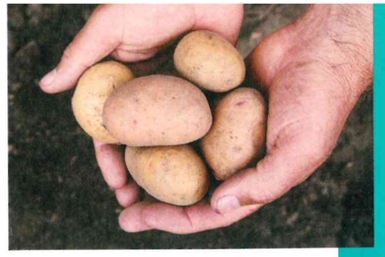
In 1964, t
moved to
unknown
DeSalvo,
was killed
many wor
the develo

Forens
contain u
DNA in ou
variable D
they come

Almost
investigat
with trace
isolate sor
only in ma
livan crim
tial match
were clear
and provi

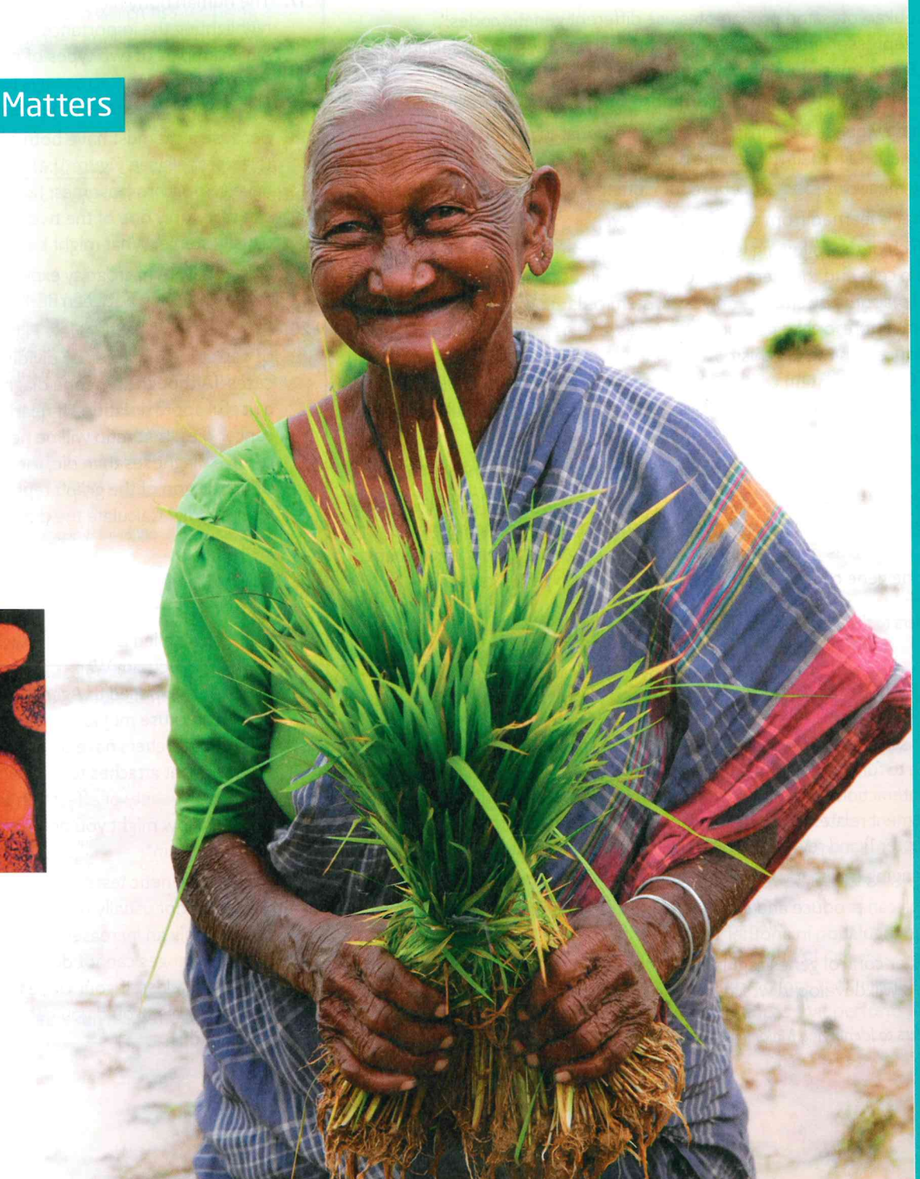
DNA te
ogy has le
produce th
shed light
detecting
technolog
social, lega

DNA Technology



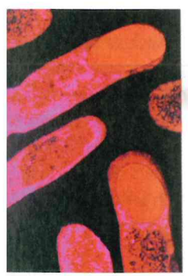
ARE MORE GENES BETTER? YOU HAVE ABOUT THE SAME NUMBER OF GENES AS A MICROSCOPIC WORM, AND ONLY HALF AS MANY AS A RICE PLANT.

SOMEDAY SOON, GENETICALLY MODIFIED POTATOES COULD SAVE TENS OF THOUSANDS OF CHILDREN FROM DEATH BY CHOLERA.



chnology Matters

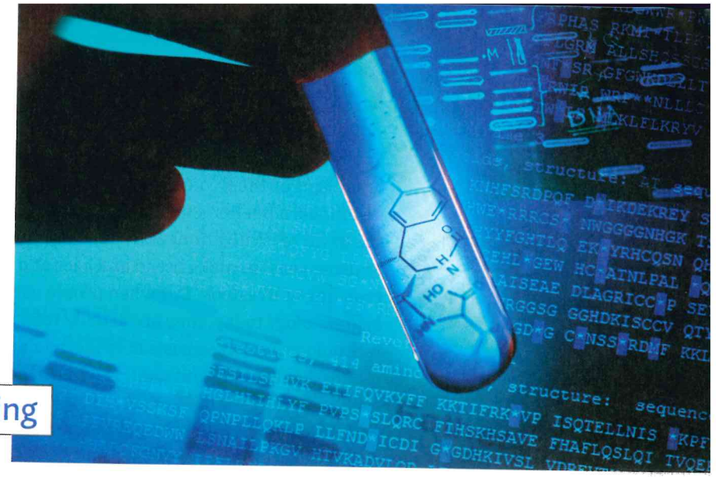
ervices
engineering
npact on
ciety. DNA
many practical
ne food we
take to other
ery day.
echnologies
evidence for
and scientific



AY:

CHAPTER THREAD

DNA Profiling



A DNA profile. Even minuscule bits of evidence can provide a DNA profile.

BIOLOGY AND SOCIETY DNA Profiling

Using DNA to Establish Guilt and Innocence

In 1964, teenager Mary Sullivan was found dead in her new apartment, just days after she had moved to Boston. She had been raped and strangled. She was one of at least 11 victims of an unknown killer nicknamed the Boston Strangler, who terrorized the city for two years. Albert DeSalvo, who had a long criminal history, confessed to the crimes, but later denied guilt. DeSalvo was killed by a fellow inmate in 1973 while in prison on separate rape convictions. For decades, many wondered whether the real Boston Strangler was still free. The answer would have to wait for the development of modern DNA technologies.

Forensic DNA analysis hinges on a simple fact: The cells of every person (except identical twins) contain unique DNA. The DNA that codes for proteins is extremely similar in all people, but other DNA in our cells does not code for anything and is highly variable from person to person. This hypervariable DNA is used for DNA profiling, which is an analysis of DNA samples to determine whether they come from the same person.

Almost 50 years after Mary Sullivan was murdered, the Boston Police Department reopened the investigation. Forensic scientists analyzed long-stored evidence from the crime, including a blanket with traces of semen. The evidence contained samples of the criminal's DNA and the victim's DNA. To isolate some of the criminal's DNA, scientists used technology that targets the Y chromosome, found only in males (and hence likely belonging to the perpetrator). The Y chromosome from the Mary Sullivan crime scene was compared with that of DeSalvo's nephew. After this comparison showed a partial match (as expected from a close relative), DeSalvo's body was exhumed for testing. The results were clear: DeSalvo was, in fact, the Boston Strangler. This conclusion brought closure to the case and provided a measure of solace to several grieving families.

DNA technology has revolutionized the field of forensics. Beyond the courtroom, DNA technology has led to some remarkable scientific advances: Crop plants have been genetically modified to produce their own insecticides; human genes are being compared with those of other animals to help shed light on what makes us distinctly human; and significant advances have been made toward detecting and curing fatal genetic diseases. This chapter will describe these and other uses of DNA technology and explain how various DNA techniques are performed. We'll also examine some of the social, legal, and ethical issues that lie at the intersection of biology and society.

Genetic Engineering

✓ CHECKPOINT

What is biotechnology? What is recombinant DNA?

■ Answer: the manipulation of organisms or their parts to produce a useful product; a molecule containing DNA from two different sources, often different species

You may think of **biotechnology**, the manipulation of organisms or their components to make useful products, as a modern phenomenon, but it actually dates back to the dawn of civilization. Consider such ancient practices as using yeast to make bread and beer and the selective breeding of livestock. But when people use the term *biotechnology* today, they are usually referring to DNA technology, which consists of modern laboratory techniques for studying and manipulating genetic material. Using the methods of DNA technology, scientists can modify specific genes and move them between organisms as different as bacteria, plants, and animals. Organisms that have acquired one or more genes by artificial means are called **genetically modified (GM) organisms**. If the newly acquired gene is from another organism, typically of another species, the recombinant organism is called a **transgenic organism**.

In the 1970s, the field of biotechnology exploded with the invention of methods for making recombinant DNA in the laboratory. Scientists construct **recombinant DNA** by combining pieces of DNA from two different sources—often from different species—in order to form a single DNA molecule. Recombinant DNA technology is widely used in **genetic engineering**, the direct manipulation of genes for practical purposes. Scientists have genetically engineered bacteria to mass-produce a variety of useful chemicals, from cancer drugs to pesticides. Scientists have also transferred genes from bacteria to plants and from one animal species to another **(Figure 12.1)**. Such engineering can serve a variety of purposes, from basic research (What does this gene do?) to medical applications (Can we create animal models for a human disease?). ✓

Recombinant DNA Techniques

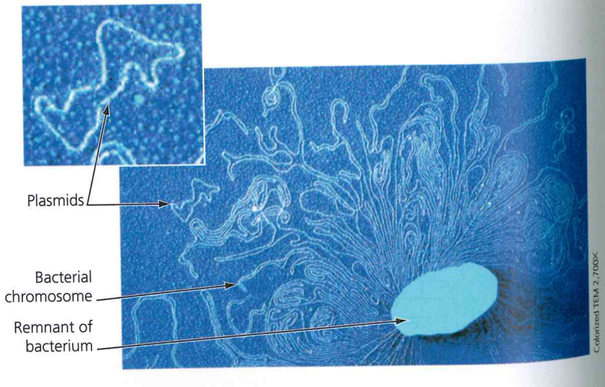
Although genetic engineering can be performed on a variety of organisms, bacteria (*Escherichia coli*, in particular) are the workhorses of modern biotechnology. To manipulate genes in the laboratory, biologists often use bacterial **plasmids**, which are small, circular DNA molecules that duplicate separately from the larger bacterial chromosome **(Figure 12.2)**. Because plasmids can carry virtually any gene and are passed from one generation of bacteria to the next, they are key tools for **DNA cloning**, the production of multiple identical copies of a piece of DNA. DNA cloning methods are central to most genetic engineering tasks.

How to Clone a Gene

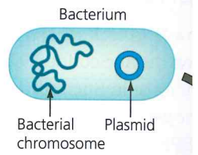
Consider a typical genetic engineering challenge: A researcher at a pharmaceutical company identifies a gene of interest that codes for a valuable protein, such as a potential new drug. The biologist wants to manufacture the protein on a large scale. **Figure 12.3** illustrates a way to accomplish this by using recombinant DNA techniques.

▼ Figure 12.1 Researchers produced glowing fish by transferring a gene for a fluorescent protein originally obtained from jellies ("jellyfish").

▼ Figure 12.2 Bacterial plasmids. The micrograph shows a bacterial cell that has been ruptured, revealing one long chromosome and several smaller plasmids. The inset is an enlarged view of a single plasmid.



To start, the biologist isolates two kinds of DNA: ba
rial plasmids that will serve as **vectors** (gene carriers)
and DNA from another organism that includes the gen
of interest (shown in yellow in the figure). This foreig
DNA may be from any type of organism, even a humar
1 The DNA from the two sources is joined, resulting i
recombinant DNA plasmids. 2 The recombinant
plasmids are then mixed with bacteria. Under the righ
conditions, the bacteria take up the recombinant plasm
3 Each bacterium, carrying its recombinant plasmid
is allowed to reproduce through cell division to form a
clone of cells, a population of genetically identical cel
descended from a single original cell. In this clone, eac
bacterium carries a copy of the gene of interest. When
DNA cloning involves a gene-carrying segment of DNA
(as it does here), it is known as **gene cloning**. As the



Reco
DNA

Recombinant bacterium

A gene for pest resistance is inserted into plants.

Some uses of genes

A gene is used to alter bacteria for cleaning up toxic waste.

Genes may be inserted into other organisms.

Genetic Engineering

You may think of **biotechnology**, the manipulation of organisms or their components to make useful products, as a modern phenomenon, but it actually dates back to the dawn of civilization. Consider such ancient practices as using yeast to make bread and beer and the selective breeding of livestock. But when people use the term *biotechnology* today, they are usually referring to DNA technology, which consists of modern laboratory techniques for studying and manipulating genetic material. Using the methods of DNA technology, scientists can modify specific genes and move them between organisms as different as bacteria, plants, and animals. Organisms that have acquired one or more genes by artificial means are called **genetically modified (GM) organisms**. If the newly acquired gene is from another organism, typically of another species, the recombinant organism is called a **transgenic organism**.

In the 1970s, the field of biotechnology exploded with the invention of methods for making recombinant DNA in the laboratory. Scientists construct **recombinant DNA** by combining pieces of DNA from two different sources—often from different species—in order to form a single DNA molecule. Recombinant DNA technology is widely used in **genetic engineering**, the direct manipulation of genes for practical purposes. Scientists have genetically engineered bacteria to mass-produce a variety of useful chemicals, from cancer drugs to pesticides. Scientists have also transferred genes from bacteria to plants and from one animal species to another **(Figure 12.1)**. Such engineering can serve a variety of purposes, from basic research (What does this gene do?) to medical applications (Can we create animal models for a human disease?). ✅

Recombinant DNA Techniques

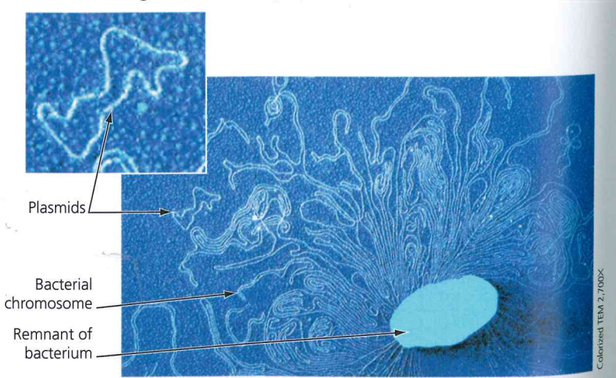
Although genetic engineering can be performed on a variety of organisms, bacteria (*Escherichia coli*, in particular) are the workhorses of modern biotechnology. To manipulate genes in the laboratory, biologists often use bacterial **plasmids**, which are small, circular DNA molecules that duplicate separately from the larger bacterial chromosome **(Figure 12.2)**. Because plasmids can carry virtually any gene and are passed from one generation of bacteria to the next, they are key tools for **DNA cloning**, the production of multiple identical copies of a piece of DNA. DNA cloning methods are central to most genetic engineering tasks.

How to Clone a Gene

Consider a typical genetic engineering challenge: A researcher at a pharmaceutical company identifies a gene of interest that codes for a valuable protein, such as a potential new drug. The biologist wants to manufacture the protein on a large scale. **Figure 12.3** illustrates a way to accomplish this by using recombinant DNA techniques.

roduced glowing fish by transferring a gene iginally obtained from jellies ("jellyfish").

▼ Figure 12.2 Bacterial plasmids. The micrograph shows a bacterial cell that has been ruptured, revealing one long chromosome and several smaller plasmids. The inset is an enlarged view of a single plasmid.



To start, the biologist isolates two kinds of DNA: bacterial plasmids that will serve as **vectors** (gene carriers) and DNA from another organism that includes the gene of interest (shown in yellow in the figure). This foreign DNA may be from any type of organism, even a human. 1 The DNA from the two sources is joined, resulting in recombinant DNA plasmids. 2 The recombinant plasmids are then mixed with bacteria. Under the right conditions, the bacteria take up the recombinant plasmids. 3 Each bacterium, carrying its recombinant plasmid, is allowed to reproduce through cell division to form a **clone** of cells, a population of genetically identical cells descended from a single original cell. In this clone, each bacterium carries a copy of the gene of interest. When DNA cloning involves a gene-carrying segment of DNA (as it does here), it is known as **gene cloning**. As the bacteria multiply, the foreign gene that is carried by the recombinant plasmid is also copied. 4 The transgenic bacteria with the gene of interest can then be grown in tanks, producing the protein in marketable quantities. The products of gene cloning may be copies of the gene itself, to be used in additional genetic engineering projects, or the protein product of the cloned gene, to be harvested and used. ✅

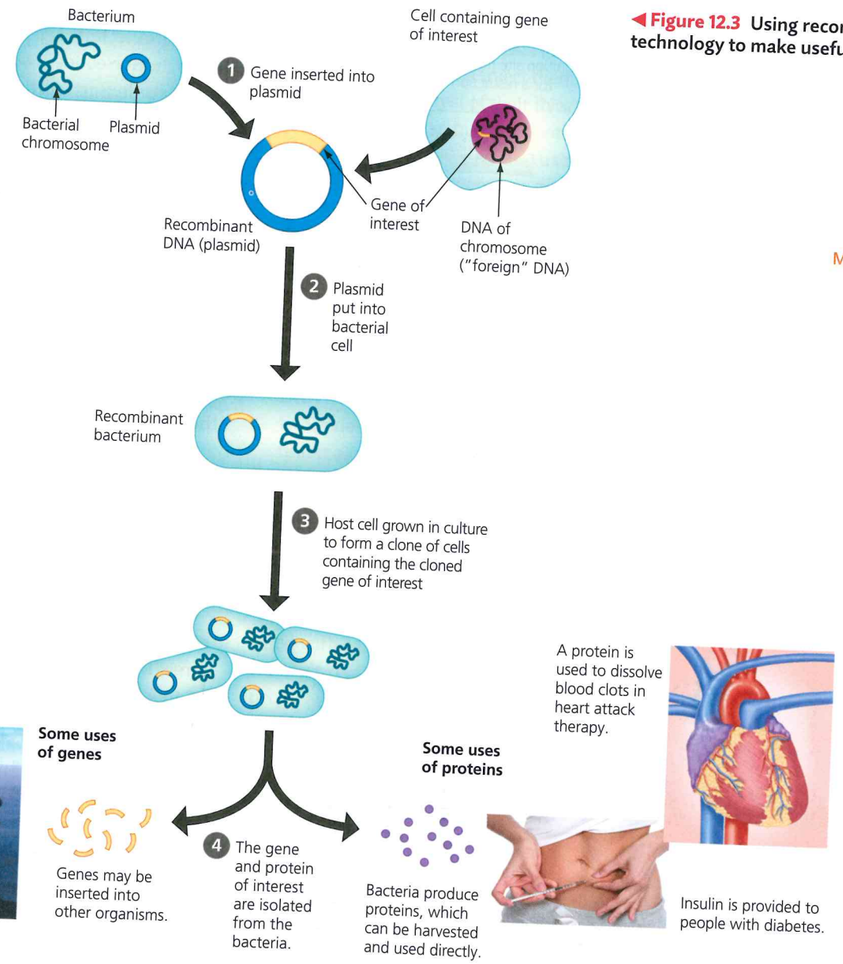
Cutting and Pasting DNA with Restriction Enzymes

As shown in Figure 12.3, recombinant DNA is created by combining two ingredients: a bacterial plasmid and the gene of interest. To understand how these DNA molecules are spliced together, you need to learn how enzymes cut and paste DNA.

✅ CHECKPOINT

Why are plasmids valuable tools for the production of recombinant DNA?

■ Answer: Plasmids can carry virtually any foreign gene and are replicated by their bacterial host cells.



◀ Figure 12.3 Using recombinant DNA technology to make useful products.

Figure Walkthrough
Mastering Biology
goo.gl/SFHdmb

✓ CHECKPOINT

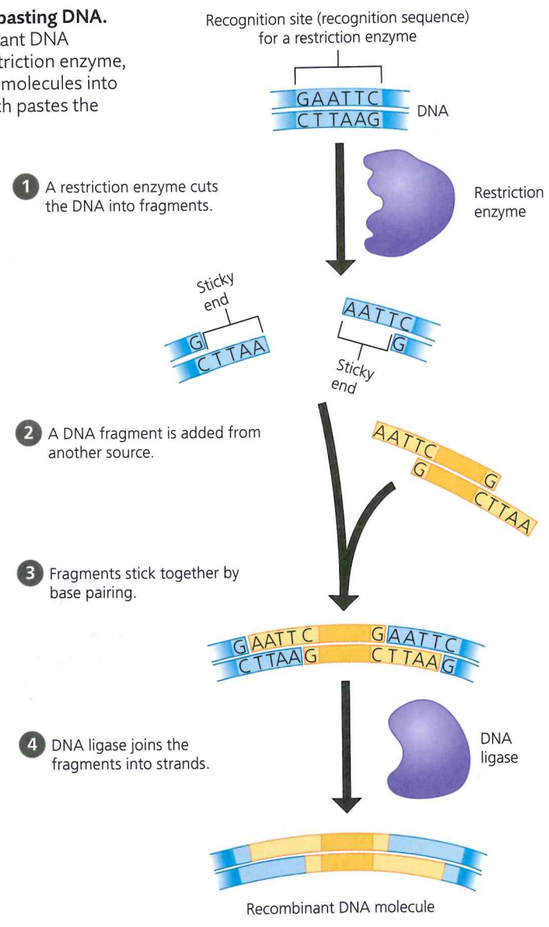
If you mix a restriction enzyme that cuts within the sequence AATC with DNA of the sequence CGAATCTAGCAATCGCGA, how many restriction fragments will result? (For simplicity, the sequence of only one of the two DNA strands is listed.)

■ *Answer: Cuts at two restriction sites will yield three restriction fragments.*

The cutting tools used for making recombinant DNA are bacterial enzymes called **restriction enzymes**. Biologists have identified hundreds of restriction enzymes, each recognizing a particular short DNA sequence, usually four to eight nucleotides long. For example, one restriction enzyme only recognizes the DNA sequence GAATTC, whereas another recognizes GGATCC. The DNA sequence recognized by a particular restriction enzyme is called a **restriction site**. After a restriction enzyme binds to its restriction site, it cuts the two strands of the DNA by breaking chemical bonds at specific points within the sequence, like a pair of highly specific molecular scissors.

The top of **Figure 12.4** shows a piece of DNA (blue) that contains one restriction site for a restriction enzyme. ❶ The restriction enzyme cuts the DNA strands between the bases A and G in the recognition sequence, producing pieces of DNA called **restriction fragments**. The staggered cuts yield two double-stranded DNA fragments with single-stranded ends, called "sticky ends." Sticky ends are the key to joining DNA restriction fragments originating from different sources. ❷ Next, a piece of DNA from another source (yellow) is added. Notice that the yellow DNA has single-stranded ends identical in base sequence to the sticky ends on the blue DNA because the same restriction enzyme was used to cut both types of DNA. ❸ The complementary ends on the blue and yellow fragments stick together by base pairing. ❹ The union between the blue and yellow fragments is then made permanent by the "pasting" enzyme **DNA ligase**. This enzyme connects the DNA pieces into continuous strands by forming bonds between adjacent nucleotides. The final outcome is a single molecule of recombinant DNA. The process just described explains what happens in step 1 of Figure 12.3. ✓

▶ Figure 12.4 Cutting and pasting DNA. The production of recombinant DNA requires two enzymes: a restriction enzyme, which cuts the original DNA molecules into pieces, and DNA ligase, which pastes the pieces together.



Gene Editing

A new DNA technology, called the **CRISPR-Cas9 system**, allows the nucleotide sequence of specific genes to be edited in living cells. Such editing can reveal the function of the gene or possibly even correct a genetic mutation. CRISPR-Cas9 is becoming one of the most important tools available for genetic engineering.

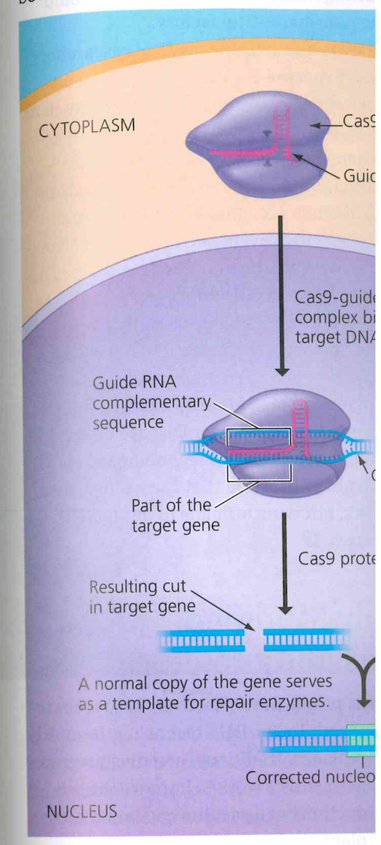
The CRISPR-Cas9 system, like restriction enzymes, was discovered as a natural component of prokaryotic cells. Scientists noticed that bacterial genomes have short repetitive DNA sequences with different stretches of "spacer DNA" between the repeats. Each spacer sequence corresponds to DNA from a virus that has infected the cell. A bacterial protein called Cas9 associated with the repeats can identify and cut viral DNA, thereby defending the bacterium against infection.

Cas9 cuts double-stranded DNA molecules (as restriction enzymes do). However, while a given restriction enzyme recognizes only one DNA sequence, the Cas9 protein will cut any DNA sequence that is complementary to an associated molecule of guide RNA. Cas9 is like a guided missile, with an RNA molecule as the guidance system.

To alter a cell, a scientist introduces a Cas9–guide RNA complex into it. The guide RNA is complementary to a target DNA sequence, such as a gene. After Cas9 cuts both strands of the target gene, DNA repair enzymes randomly insert nucleotides as they reconnect the target DNA. In this way, researchers can "knock out" (disable) a given gene. By observing the altered cell or organism, researchers may be able to determine the function of the knocked-out gene.

The CRISPR-Cas9 system can also be used to edit a gene **(Figure 12.5)**. Researchers can introduce a segment from the normal gene along with the Cas9–guide RNA complex. After Cas9 cuts the target DNA, repair enzymes use the normal DNA as a template to repair the target

▼ Figure 12.5 The CRISPR-Cas9 system. By
RNA "homing" sequence with an enzyme, spe
be inactivated or editing within living cells.



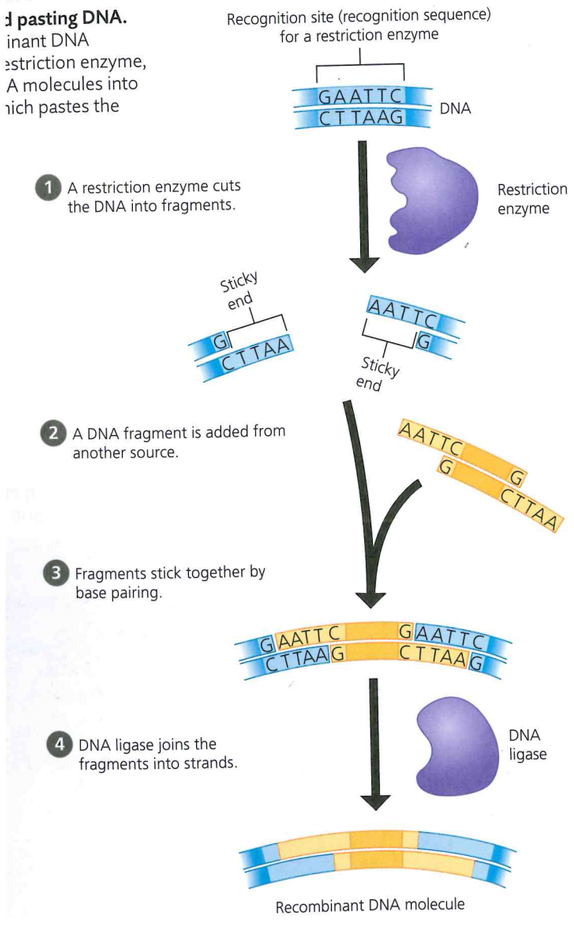
DNA where it was cut. In this way, the CR
tem acts like the "search and replace" fur
processor, potentially fixing mutations ir

In 2014, researchers used the CRISPR-C
fix a genetic defect in mice. They altered li
to correct a faulty gene that causes tyrosin
affecting metabolism of the amino acid tyr
lead to organ dysfunction and developmer
A study in 2015 involved mice that carriec
gene that codes for dystrophin, a protein e
cle function. Researchers infected the mic
rying the Cas9–guide RNA complex. The v
muscle cells and removed a region of the g
the dystrophin mutation. The gene, now la
tion, produced normal dystrophin protein
the muscles to function properly. Researcl
this technique to humans carrying a simil
causes the disease Duchenne muscular dy

The cutting tools used for making recombinant DNA are bacterial enzymes called **restriction enzymes**. Biologists have identified hundreds of restriction enzymes, each recognizing a particular short DNA sequence, usually four to eight nucleotides long. For example, one restriction enzyme only recognizes the DNA sequence GAATTC, whereas another recognizes GGATCC. The DNA sequence recognized by a particular restriction enzyme is called a **restriction site**. After a restriction enzyme binds to its restriction site, it cuts the two strands of the DNA by breaking chemical bonds at specific points within the sequence, like a pair of highly specific molecular scissors.

The top of **Figure 12.4** shows a piece of DNA (blue) that contains one restriction site for a restriction enzyme. 1 The restriction enzyme cuts the DNA strands between the bases A and G in the recognition sequence, producing pieces of DNA called **restriction fragments**. The staggered cuts yield two double-stranded DNA fragments with single-stranded ends, called "sticky ends." Sticky ends are the key to joining DNA restriction fragments originating from different sources. 2 Next, a piece of DNA from another source (yellow) is added. Notice that the yellow DNA has single-stranded ends identical in base sequence to the sticky ends on the blue DNA because the same restriction enzyme was used to cut both types of DNA. 3 The complementary ends on the blue and yellow fragments stick together by base pairing. 4 The union between the blue and yellow fragments is then made permanent by the "pasting" enzyme **DNA ligase**. This enzyme connects the DNA pieces into continuous strands by forming bonds between adjacent nucleotides. The final outcome is a single molecule of recombinant DNA. The process just described explains what happens in step 1 of Figure 12.3. ✅

d pasting DNA.
inant DNA
estriction enzyme,
A molecules into
hich pastes the



Gene Editing

A new DNA technology, called the **CRISPR-Cas9 system**, allows the nucleotide sequence of specific genes to be edited in living cells. Such editing can reveal the function of the gene or possibly even correct a genetic mutation. CRISPR-Cas9 is becoming one of the most important tools available for genetic engineering.

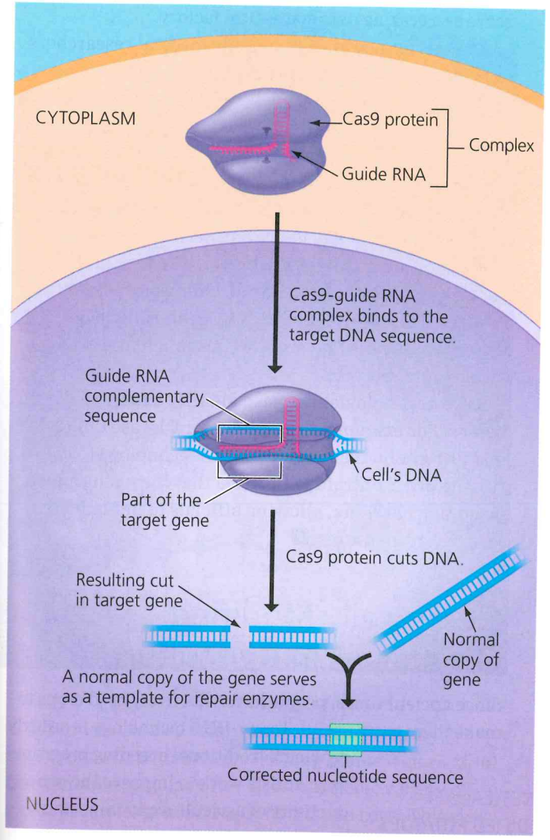
The CRISPR-Cas9 system, like restriction enzymes, was discovered as a natural component of prokaryotic cells. Scientists noticed that bacterial genomes have short repetitive DNA sequences with different stretches of "spacer DNA" between the repeats. Each spacer sequence corresponds to DNA from a virus that has infected the cell. A bacterial protein called Cas9 associated with the repeats can identify and cut viral DNA, thereby defending the bacterium against infection.

Cas9 cuts double-stranded DNA molecules (as restriction enzymes do). However, while a given restriction enzyme recognizes only one DNA sequence, the Cas9 protein will cut any DNA sequence that is complementary to an associated molecule of guide RNA. Cas9 is like a guided missile, with an RNA molecule as the guidance system.

To alter a cell, a scientist introduces a Cas9–guide RNA complex into it. The guide RNA is complementary to a target DNA sequence, such as a gene. After Cas9 cuts both strands of the target gene, DNA repair enzymes randomly insert nucleotides as they reconnect the target DNA. In this way, researchers can "knock out" (disable) a given gene. By observing the altered cell or organism, researchers may be able to determine the function of the knocked-out gene.

The CRISPR-Cas9 system can also be used to edit a gene **(Figure 12.5)**. Researchers can introduce a segment from the normal gene along with the Cas9–guide RNA complex. After Cas9 cuts the target DNA, repair enzymes use the normal DNA as a template to repair the target DNA where it was cut. In this way, the CRISPR-Cas9 system acts like the "search and replace" function of a word processor, potentially fixing mutations in cells.

▼ **Figure 12.5 The CRISPR-Cas9 system.** By combining an RNA "homing" sequence with an enzyme, specific genes can be inactivated or editing within living cells.



In 2014, researchers used the CRISPR-Cas9 system to fix a genetic defect in mice. They altered live mouse cells to correct a faulty gene that causes tyrosinemia, a disease affecting metabolism of the amino acid tyrosine, which can lead to organ dysfunction and developmental disabilities. A study in 2015 involved mice that carried a mutation in a gene that codes for dystrophin, a protein essential for muscle function. Researchers infected the mice with a virus carrying the Cas9–guide RNA complex. The virus infected the muscle cells and removed a region of the gene containing the dystrophin mutation. The gene, now lacking the mutation, produced normal dystrophin proteins, which allowed the muscles to function properly. Researchers hope to apply this technique to humans carrying a similar mutation that causes the disease Duchenne muscular dystrophy.

There are many hurdles to clear before the CRISPR-Cas9 system can be tried in humans, but the technique is sparking the interest of researchers and physicians all around the world. In the next section, we'll explore some ways that genetic modifications affect your life. ✅

Medical Applications

By transferring the gene for a desired protein into a bacterium, yeast, or other kind of cell that is easy to grow in culture, scientists can produce large quantities of useful proteins that are present naturally only in small amounts. In this section, you'll learn about some applications of recombinant DNA technology.

Humulin is human insulin produced by genetically modified bacteria **(Figure 12.6)**. In humans, insulin is a protein normally made by the pancreas. Insulin functions as a hormone and helps regulate the level of glucose in the blood. If the body fails to produce enough insulin, the result is type 1 diabetes. There is no cure, so people with this disease must inject themselves daily with doses of insulin for the rest of their lives.

Because human insulin is not readily available, diabetes was historically treated using insulin from cows and pigs. This treatment was problematic, however. Pig and cow insulins can cause allergic reactions in people because their chemical structures differ slightly from that of human insulin. In addition, by the 1970s, the supply of beef and pork pancreas available for insulin extraction could not keep up with the demand.

In 1978, scientists working at a biotechnology company chemically synthesized DNA fragments and linked them to form the two genes that code for the two polypeptides that make up human insulin. They then inserted these artificial genes into *E. coli* host cells. Under proper growing conditions, the transgenic bacteria cranked out large quantities of the human protein. In 1982, Humulin hit the market as the world's first genetically engineered pharmaceutical product. Today, it is produced around the clock in gigantic fermentation vats filled with a liquid culture of bacteria. Each day, more than 4 million people with diabetes use the insulin collected, purified, and packaged at such facilities **(Figure 12.7)**.

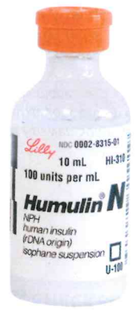
Insulin is just one of many human proteins produced by genetically modified bacteria. Another example is human growth hormone (HGH). Abnormally low levels of this hormone during childhood and adolescence can cause dwarfism.

✅ CHECKPOINT

What is the function of the guide RNA in the CRISPR-Cas9 system?

■ Answer: It guides the complex to the proper location in the genome.

▼ **Figure 12.6 Humulin, human insulin produced by genetically modified bacteria.**



MILLIONS OF PEOPLE WITH DIABETES LIVE HEALTHIER LIVES THANKS TO INSULIN MADE BY BACTERIA.

▼ **Figure 12.7 A factory that produces genetically engineered insulin.**

Because growth hormones from other animals are not effective in people, HGH was an early target of genetic engineers. Before genetically engineered HGH became available in 1985, children with an HGH deficiency could only be treated with scarce and expensive supplies of HGH obtained from human cadavers. Another genetically engineered protein helps dissolve blood clots. If administered shortly after a stroke, it reduces the risk of additional strokes and heart attacks.

Although bacteria can produce many human proteins, some proteins can only be made by eukaryotic cells, such as cells from fungi, animals, and plants. Common baker's yeast is currently used to produce proteins used as medicines, including the hepatitis B vaccine, an antimalarial drug, and interferons used to treat cancer and viral infections. In 2015, scientists announced they had transferred 23 genes (from bacteria, plants, and animals) into yeast that allow the recombinant fungi to convert sugar into the painkiller drug hydrocodone. Genetically modified mammalian cells are used to produce erythropoietin (EPO), a hormone used to treat anemia by stimulating production of red blood cells. Researchers have also developed transgenic plant cells that can produce human drugs. The drug factories of the future may be carrots because they are easily grown in culture and are unlikely to be contaminated by human pathogens (such as viruses).

Genetically modified whole animals are also used to produce drugs. **Figure 12.8** shows a transgenic goat that carries a gene for an enzyme called lysozyme. This enzyme, found naturally in breast milk, has antibacterial properties. In another example, the gene for a human blood protein has been inserted into the genome of a goat so that the protein is secreted in the goat's milk. The protein is then purified from the milk. Because transgenic animals are difficult to produce, researchers may create a single transgenic animal and then breed or clone it. The resulting herd of transgenic animals could serve as a grazing pharmaceutical factory.

DNA technology is also helping medical researchers develop vaccines. A vaccine is a harmless variant or derivative of a disease-causing microorganism—such as a bacterium or virus—that is used to prevent an infectious disease. When a person is inoculated, the vaccine stimulates the immune system to develop lasting defenses against the microorganism. For many viral diseases, the only way to prevent serious harm is to use vaccination to prevent the illness in the first place. The vaccine against hepatitis B, a disabling and sometimes fatal liver disease, is produced by genetically engineered yeast cells that secrete a protein found on the virus's outer surface.

DNA technologies can also identify causes of illnesses. For example, the Centers for Disease Control and Prevention regularly uses DNA technology to identify the precise strain of bacteria that is causing a food poisoning outbreak, allowing officials to implement food safety measures. ✅

✅ CHECKPOINT

Bacteria are easier to manipulate than yeast. So why are yeast used to produce some human proteins?

■ Answer: Some genes require eukaryotic cells in order to be properly made.

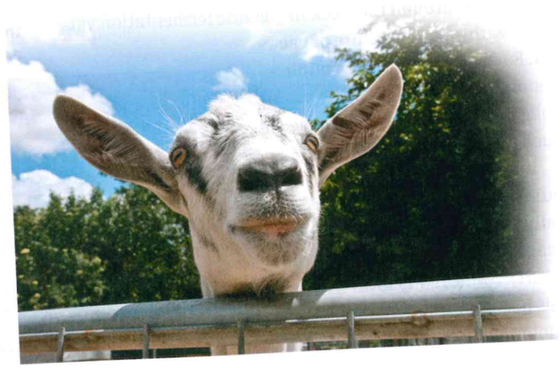
GENETICALLY MODIFIED POTATOES COULD SAVE MANY CHILDREN FROM DEATH BY CHOLERA.

Genetically Modified Organisms in Agriculture

Since ancient times, people have selectively bred crops to make them more useful. Today, DNA technology is quickly replacing traditional breeding programs as scientists work to improve the productivity of agriculturally important plants and animals.

In the United States today, nearly all of our corn, soybean, and cotton crops are genetically modified. **Figure 12.9** shows corn that has been genetically engineered to resist attack by an insect called the European corn borer. Growing insect-resistant plants reduces the need for chemical insecticides. In another example, modified strawberry plants produce bacterial proteins that act as a natural antifreeze, protecting the delicate plants from the damages of cold weather. Potatoes and rice have been engineered to produce harmless proteins derived from the cholera bacterium; researchers hope that these modified foods will one day serve as an edible vaccine against cholera, a disease that kills thousands of children in developing nations every year. In India, the insertion of a natural but rare saltwater resistance gene has enabled new varieties of rice to thrive in water three times as salty as seawater, allowing food to be grown in drought-stricken or flooded regions.

▶ **Figure 12.8** **A genetically modified goat.**



▼ **Figure 12.9** **Genetically modified corn.** The co this field carry a bacterial gene that helps prevent by the European corn borer (inset).

Scientists are also using genetic engineerin improve the nutritional value of crop plants (One example is "golden rice 2," a transgenic v rice that carries genes from daffodils and corr

▼ **Figure 12.10** **Genetically modified stapl** ordinary rice, has been genetically modified converts to vitamin A. Transgenic cassava (r nearly a billion people, has been modified t

Because growth hormones from other animals are not effective in people, HGH was an early target of genetic engineers. Before genetically engineered HGH became available in 1985, children with an HGH deficiency could only be treated with scarce and expensive supplies of HGH obtained from human cadavers. Another genetically engineered protein helps dissolve blood clots. If administered shortly after a stroke, it reduces the risk of additional strokes and heart attacks.

Although bacteria can produce many human proteins, some proteins can only be made by eukaryotic cells, such as cells from fungi, animals, and plants. Common baker's yeast is currently used to produce proteins used as medicines, including the hepatitis B vaccine, an antimalarial drug, and interferons used to treat cancer and viral infections. In 2015, scientists announced they had transferred 23 genes (from bacteria, plants, and animals) into yeast that allow the recombinant fungi to convert sugar into the painkiller drug hydrocodone. Genetically modified mammalian cells are used to produce erythropoietin (EPO), a hormone used to treat anemia by stimulating production of red blood cells. Researchers have also developed transgenic plant cells that can produce human drugs. The drug factories of the future may be carrots because they are easily grown in culture and are unlikely to be contaminated by human pathogens (such as viruses).

Genetically modified whole animals are also used to produce drugs. **Figure 12.8** shows a transgenic goat that carries a gene for an enzyme called lysozyme. This enzyme, found naturally in breast milk, has antibacterial properties. In another example, the gene for a human blood protein has been inserted into the genome of a goat so that the protein is secreted in the goat's milk. The protein is then purified from the milk. Because transgenic animals are difficult to produce, researchers may create a single transgenic animal and then breed or clone it. The resulting herd of transgenic animals could serve as a grazing pharmaceutical factory.

GENETICALLY MODIFIED POTATOES COULD SAVE MANY CHILDREN FROM DEATH BY CHOLERA.

DNA technology is also helping medical researchers develop vaccines. A vaccine is a harmless variant or derivative of a disease-causing microorganism—such as a bacterium or virus—that is used to prevent an infectious disease. When a person is inoculated, the vaccine stimulates the immune system to develop lasting defenses against the microorganism. For many viral diseases, the only way to prevent serious harm is to use vaccination to prevent the illness in the first place. The vaccine against hepatitis B, a disabling and sometimes fatal liver disease, is produced by genetically engineered yeast cells that secrete a protein found on the virus's outer surface.

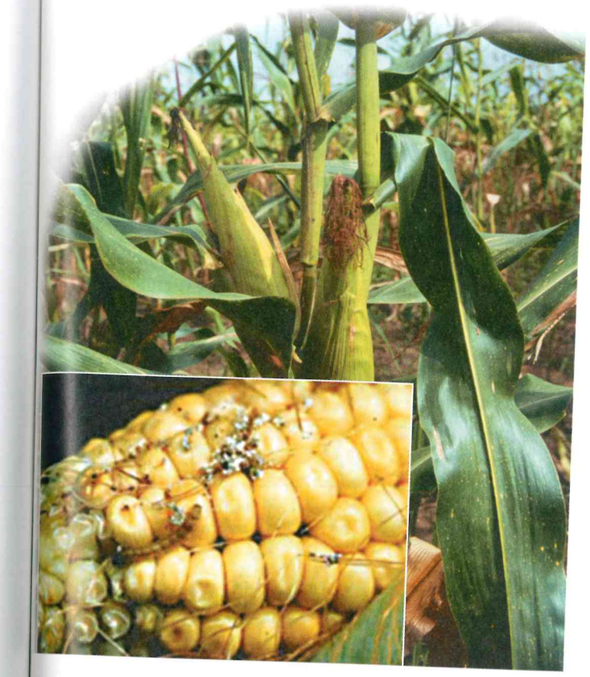
DNA technologies can also identify causes of illnesses. For example, the Centers for Disease Control and Prevention regularly uses DNA technology to identify the precise strain of bacteria that is causing a food poisoning outbreak, allowing officials to implement food safety measures. ✅

Genetically Modified Organisms in Agriculture

Since ancient times, people have selectively bred crops to make them more useful. Today, DNA technology is quickly replacing traditional breeding programs as scientists work to improve the productivity of agriculturally important plants and animals.

In the United States today, nearly all of our corn, soybean, and cotton crops are genetically modified. **Figure 12.9** shows corn that has been genetically engineered to resist attack by an insect called the European corn borer. Growing insect-resistant plants reduces the need for chemical insecticides. In another example, modified strawberry plants produce bacterial proteins that act as a natural antifreeze, protecting the delicate plants from the damages of cold weather. Potatoes and rice have been engineered to produce harmless proteins derived from the cholera bacterium; researchers hope that these modified foods will one day serve as an edible vaccine against cholera, a disease that kills thousands of children in developing nations every year. In India, the insertion of a natural but rare saltwater resistance gene has enabled new varieties of rice to thrive in water three times as salty as seawater, allowing food to be grown in drought-stricken or flooded regions.

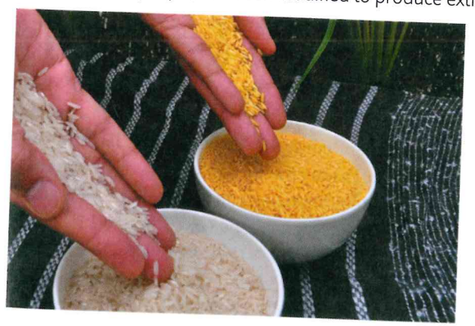
▼ **Figure 12.9 Genetically modified corn.** The corn plants in this field carry a bacterial gene that helps prevent infestation by the European corn borer (inset).



Scientists are also using genetic engineering to improve the nutritional value of crop plants **(Figure 12.10)**. One example is "golden rice 2," a transgenic variety of rice that carries genes from daffodils and corn. This rice could help prevent vitamin A deficiency and resulting blindness, especially in developing nations that depend on rice as a staple crop. Cassava, a starchy root crop that is a staple for nearly 1 billion people in developing nations, has similarly been modified to produce increased levels of iron and beta-carotene (which is converted to vitamin A in the body). However, controversy surrounds the use of GM foods, as we'll discuss at the end of the chapter.

Genetic engineers are targeting agricultural animals as well as plant crops. Scientists might, for example, identify in one variety of cattle a gene that causes the development of larger muscles (which make up most of the meat we eat) and transfer it to other cattle or even to chickens. Researchers have genetically modified pigs to carry a roundworm gene whose protein converts less healthy fatty acids to omega-3 fatty acids. Meat from the modified pigs contains four to five times as much healthy omega-3 fat as regular pork. In 2015, researchers replaced a gene in dairy cows with one from Angus cattle to produce cattle that lack horns, saving the bulls from painful dehorning. Similar gene-editing techniques produce improved varieties of goats (for meat and cashmere wool), pigs (for agriculture and pets), and dogs. A type of Atlantic salmon has been genetically modified to reach market size in half the normal time (18 months versus 3 years) and to grow twice as large. In late 2015, the FDA approved the sale of this GMO salmon to U.S. consumers, declaring that it is as safe and nutritious as traditional salmon. Although it could be years before the GMO salmon reaches store shelves, this is the first time a transgenic animal product was allowed to be sold as food in the United States. ✅

▼ **Figure 12.10 Genetically modified staple crops.** "Golden rice 2," the yellow grains shown here (left) alongside ordinary rice, has been genetically modified to produce high levels of beta-carotene, a molecule that the body converts to vitamin A. Transgenic cassava (right), a starchy root crop that serves as the main food source for nearly a billion people, has been modified to produce extra nutrients.



✅ CHECKPOINT

What is a genetically modified organism?

■ *Answer: one that carries DNA introduced through artificial means*

Human Gene Therapy

We've seen that bacteria, fungi, plants, and nonhuman animals can be genetically modified—so what about humans? **Human gene therapy** is intended to treat disease by introducing new genes into an afflicted person. In some cases, a mutant version of a gene may be replaced or supplemented with the normal allele, potentially correcting a genetic disorder, perhaps permanently. In other cases, genes are inserted and expressed only long enough to treat a medical problem.

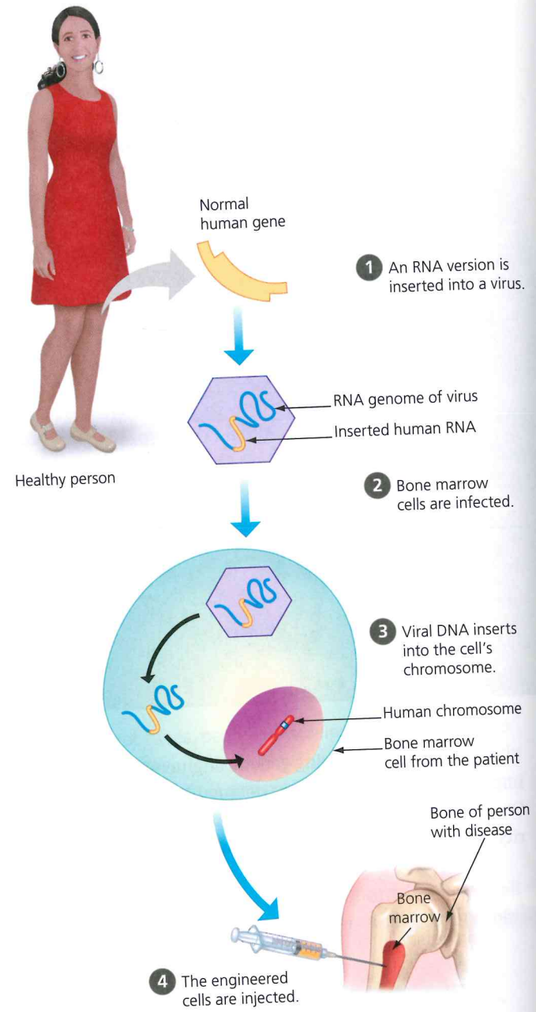
Figure 12.11 summarizes one approach to human gene therapy. The procedure closely resembles the gene cloning process shown in steps 1 through 3 of Figure 12.3, but in this instance human cells, rather than bacteria, are the targets. (1) A gene from a normal person is cloned, converted to an RNA version, and then inserted into the RNA genome of a harmless virus. (2) Bone marrow cells are taken from the patient and infected with the recombinant virus. (3) The virus inserts a DNA copy of its genome, including the normal human gene, into the DNA of the patient's cells. (4) The engineered cells are then injected back into the patient. The normal gene is transcribed and translated within the patient's body, producing the desired protein. Ideally, the nonmutant version of the gene would be inserted into cells that multiply throughout a person's life. Bone marrow cells, which include the stem cells that give rise to all the types of blood cells, are prime candidates. If the procedure succeeds, the cells will multiply permanently and produce a steady supply of the missing protein, curing the patient.

The promise of gene therapy exceeds actual results, but there have been some successes. In 2009, an international research team conducted a trial that focused on a form of progressive blindness linked to a defect in a gene responsible for producing light-detecting pigments in the eye. The researchers found that an injection of a virus carrying the normal gene into one eye of affected children improved vision in that eye, sometimes enough to allow normal function, without significant side effects. The other eye was left untreated as a control.

From 2000 to 2011, gene therapy cured 22 children with severe combined immunodeficiency (SCID), a fatal inherited disease caused by a defective gene that prevents development of the immune system, requiring patients to remain isolated within protective "bubbles." Unless treated with a bone marrow transplant, which is effective only 60% of the time, SCID patients quickly die from infections that most of us easily fend off. In these cases, researchers periodically removed immune system cells from the patients' blood, infected them with a virus engineered to carry the normal allele of the defective gene, then reinjected the blood into the patient. The treatment cured the patients of SCID, but there were serious side effects: Four of the treated patients developed leukemia, and one died after the inserted gene activated an oncogene (see Chapter 11), creating cancerous blood cells.

Two other illnesses that have been treated with gene therapy, if only in a few patients, are a degenerative disease of the nervous system and a blood disorder involving a hemoglobin gene. Research on gene therapy continues, with tougher guidelines for safe and effective application. ☑

▼ **Figure 12.11** One approach to human gene therapy.



☑ **CHECKPOINT**

Why are bone marrow stem cells ideally suited as targets for gene therapy?

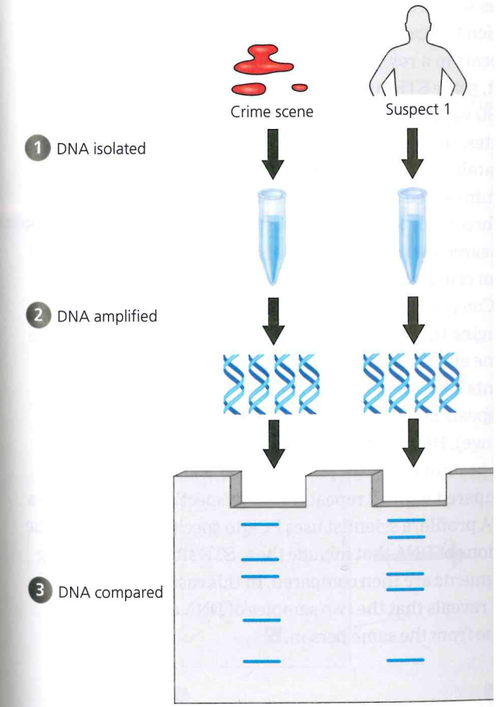
■ Answer: because bone marrow stem cells multiply throughout a person's life

DNA Profiling and F

When a crime is committed, body fluids (such as blood or semen) or small pieces of tissue (such as skin beneath a vi tim's fingernails) may be left at the scene or on the victim assailant. As discussed in the Biology and Society section, such evidence can be examined by **DNA profiling**, the analysis of DNA samples to determine whether they com from the same individual. Indeed, DNA profiling has rapidly transformed the field of **forensics**, the scientific ana sis of evidence for crime scene investigations and other le proceedings. To produce a DNA profile, scientists compar DNA sequences that vary from person to person.

Figure 12.12 presents an overview of a typical investigation using DNA profiling. (1) First, DNA samples are isolated from the crime scene, suspects, victims, or othe evidence. (2) Next, selected sequences from each DNA sample are amplified (copied many times) to produce a large sample of DNA fragments. (3) Finally, the amplified DNA fragments are compared. Together, these step provide data about which samples are from the same in vidual and which samples are unique.

▼ **Figure 12.12 Overview of DNA profiling.** In this example DNA from suspect 1 does not match DNA found at the crim scene, but DNA from suspect 2 does match.



Human Gene Therapy

We've seen that bacteria, fungi, plants, and nonhuman animals can be genetically modified—so what about humans? **Human gene therapy** is intended to treat disease by introducing new genes into an afflicted person. In some cases, a mutant version of a gene may be replaced or supplemented with the normal allele, potentially correcting a genetic disorder, perhaps permanently. In other cases, genes are inserted and expressed only long enough to treat a medical problem.

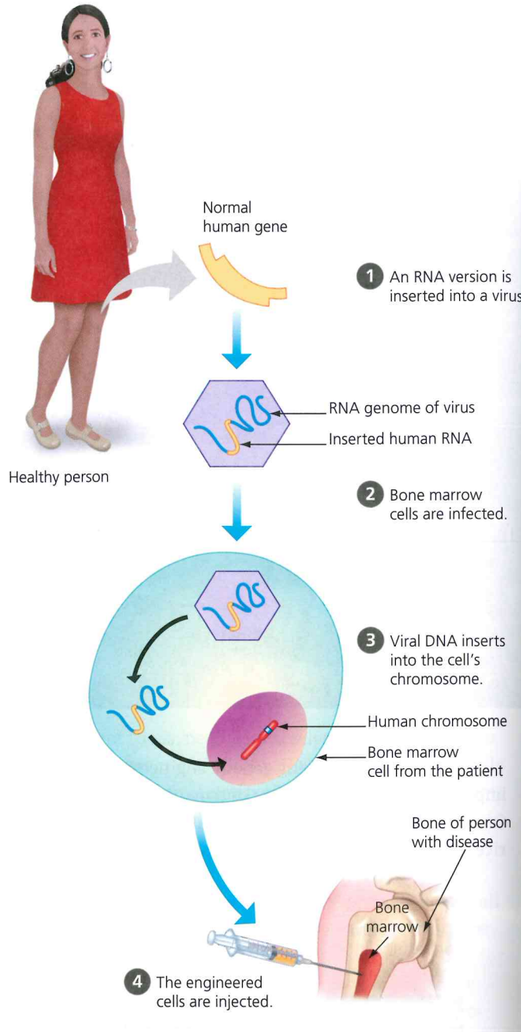
Figure 12.11 summarizes one approach to human gene therapy. The procedure closely resembles the gene cloning process shown in steps 1 through 3 of Figure 12.3, but in this instance human cells, rather than bacteria, are the targets. 1 A gene from a normal person is cloned, converted to an RNA version, and then inserted into the RNA genome of a harmless virus. 2 Bone marrow cells are taken from the patient and infected with the recombinant virus. 3 The virus inserts a DNA copy of its genome, including the normal human gene, into the DNA of the patient's cells. 4 The engineered cells are then injected back into the patient. The normal gene is transcribed and translated within the patient's body, producing the desired protein. Ideally, the nonmutant version of the gene would be inserted into cells that multiply throughout a person's life. Bone marrow cells, which include the stem cells that give rise to all the types of blood cells, are prime candidates. If the procedure succeeds, the cells will multiply permanently and produce a steady supply of the missing protein, curing the patient.

The promise of gene therapy exceeds actual results, but there have been some successes. In 2009, an international research team conducted a trial that focused on a form of progressive blindness linked to a defect in a gene responsible for producing light-detecting pigments in the eye. The researchers found that an injection of a virus carrying the normal gene into one eye of affected children improved vision in that eye, sometimes enough to allow normal function, without significant side effects. The other eye was left untreated as a control.

From 2000 to 2011, gene therapy cured 22 children with severe combined immunodeficiency (SCID), a fatal inherited disease caused by a defective gene that prevents development of the immune system, requiring patients to remain isolated within protective "bubbles." Unless treated with a bone marrow transplant, which is effective only 60% of the time, SCID patients quickly die from infections that most of us easily fend off. In these cases, researchers periodically removed immune system cells from the patients' blood, infected them with a virus engineered to carry the normal allele of the defective gene, then reinjected the blood into the patient. The treatment cured the patients of SCID, but there were serious side effects: Four of the treated patients developed leukemia, and one died after the inserted gene activated an oncogene (see Chapter 11), creating cancerous blood cells.

Two other illnesses that have been treated with gene therapy, if only in a few patients, are a degenerative disease of the nervous system and a blood disorder involving a hemoglobin gene. Research on gene therapy continues, with tougher guidelines for safe and effective application. ✅

▼ **Figure 12.11** One approach to human gene therapy.

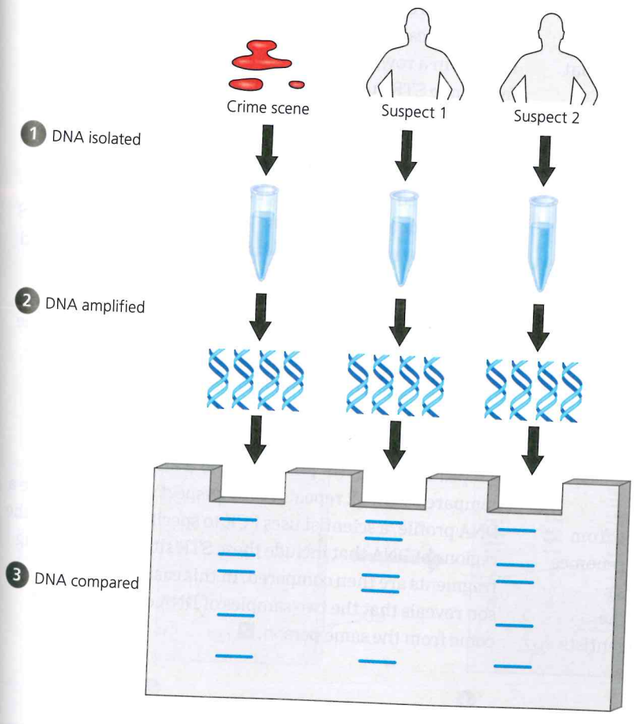


DNA Profiling and Forensic Science

When a crime is committed, body fluids (such as blood or semen) or small pieces of tissue (such as skin beneath a victim's fingernails) may be left at the scene or on the victim or assailant. As discussed in the Biology and Society section, such evidence can be examined by **DNA profiling**, the analysis of DNA samples to determine whether they come from the same individual. Indeed, DNA profiling has rapidly transformed the field of **forensics**, the scientific analysis of evidence for crime scene investigations and other legal proceedings. To produce a DNA profile, scientists compare DNA sequences that vary from person to person.

Figure 12.12 presents an overview of a typical investigation using DNA profiling. 1 First, DNA samples are isolated from the crime scene, suspects, victims, or other evidence. 2 Next, selected sequences from each DNA sample are amplified (copied many times) to produce a large sample of DNA fragments. 3 Finally, the amplified DNA fragments are compared. Together, these steps provide data about which samples are from the same individual and which samples are unique.

▼ **Figure 12.12** **Overview of DNA profiling.** In this example, DNA from suspect 1 does not match DNA found at the crime scene, but DNA from suspect 2 does match.



DNA Profiling Techniques

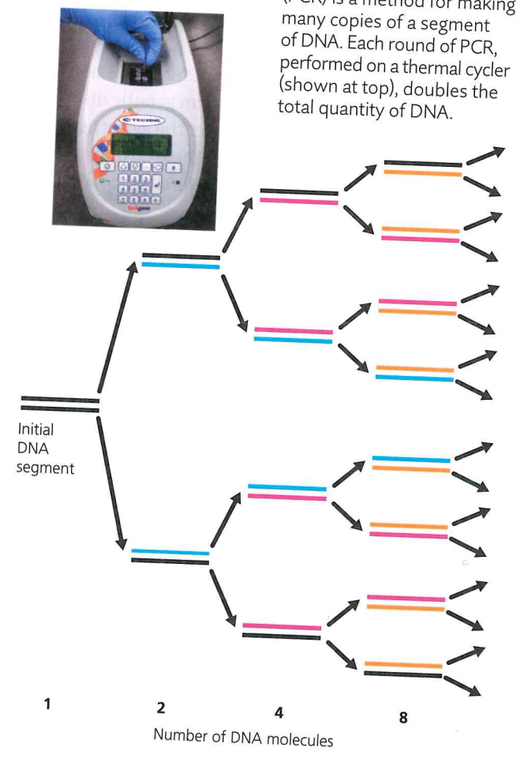
In this section, we'll look at how a DNA profile is made. Researchers use three basic techniques: the polymerase chain reaction, short tandem repeat analysis, and gel electrophoresis.

The Polymerase Chain Reaction (PCR)

The **polymerase chain reaction (PCR)** is a technique by which a specific segment of DNA can be amplified—that is, targeted and copied quickly and precisely. Through PCR, a scientist can obtain enough DNA from even minute amounts of blood or other tissue to allow a DNA profile to be constructed. In fact, a microscopic sample with as few as 20 cells can be sufficient for PCR amplification.

In principle, PCR is simple. A DNA sample is mixed with nucleotides, the DNA replication enzyme DNA polymerase, and a few other ingredients. The solution is then exposed to cycles of heating (to separate the DNA strands) and cooling (to allow double-stranded DNA to re-form). During these cycles, specific regions of each molecule of DNA are replicated, doubling the amount of that DNA (Figure 12.13). The result of this chain reaction

▼ **Figure 12.13** **DNA amplification by PCR.** The polymerase chain reaction (PCR) is a method for making many copies of a segment of DNA. Each round of PCR, performed on a thermal cycler (shown at top), doubles the total quantity of DNA.



is an exponentially growing population of identical DNA molecules. The key to automated PCR is an unusually heat-stable DNA polymerase, first isolated from prokaryotes living in hot springs (see Figure 15.16). Unlike most proteins—which denature, or fall apart, at high temperatures—this enzyme can withstand the heat at the start of each cycle.

DNA molecules are typically very long, but usually only a very small target region of a large DNA molecule needs to be amplified. The key to amplifying one particular segment of DNA and no others is the use of **primers**, short (usually 15–20 nucleotides long), chemically synthesized single-stranded DNA molecules. For each experiment, specific primers are chosen that are complementary to sequences found only at each end of the target sequence. The primers thus bind to sequences that flank the target sequence, marking the start and end points for the segment of DNA to be amplified. Beginning with a single DNA molecule and the appropriate primers, automated PCR can generate hundreds of billions of copies of the desired sequence in a few hours.

In addition to forensic applications, PCR can be used in the treatment and diagnosis of disease. For example, because the sequence of the genome of HIV (the virus that causes AIDS) is known, PCR can be used to amplify, and thus detect, HIV in blood or tissue samples. In fact, PCR is often the best way to detect this otherwise elusive virus. Medical scientists can now diagnose hundreds of human genetic disorders by using PCR with primers that target the genes associated with these disorders. The amplified DNA product is then studied to reveal the presence or absence of the disease-causing mutation. Among the genes for human diseases that have been identified are those for sickle-cell disease, hemophilia, cystic fibrosis, Huntington's disease, and Duchenne muscular dystrophy. Individuals afflicted with such diseases can often be identified before the onset of symptoms, even before birth, allowing for preventative medical care to begin. PCR can also be used to identify symptomless carriers of potentially harmful recessive alleles (see Figure 9.14). Parents may thus be informed of whether they have a risk of bearing a child with a rare disease that they do not themselves display.

☑ CHECKPOINT

1. Why is only the slightest trace of DNA at a crime scene often sufficient for forensic analysis?
2. What are STRs, and why are they useful for DNA profiling?

■ Answers: **1.** because PCR can be used to produce enough molecules for analysis **2.** STRs are nucleotide sequences repeated many times in a row. STRs are valuable for DNA profiling because different people have different numbers of repeats at the various STR sites.

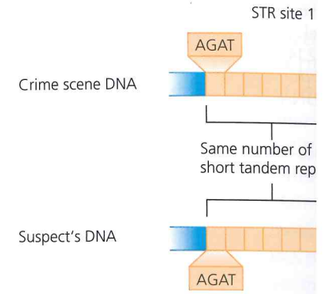
Short Tandem Repeat (STR) Analysis

How do you prove that two samples of DNA come from the same person? You could compare the entire genomes found in the two samples. But such an approach is impractical because the DNA of two humans of the same sex is 99.9% identical. Instead, forensic scientists typically compare about a dozen short segments of noncoding repetitive DNA that are known to vary from person to person. Have you ever seen a puzzle in a magazine that presents two nearly identical photos and asks you to find the few differences between them? In a similar way, scientists can focus on the few areas of difference in the human genome, ignoring the identical majority.

Repetitive DNA, which makes up much of the DNA that lies between genes in humans, consists of nucleotide sequences that are present in multiple copies in the genome. Some of this DNA consists of short sequences repeated many times tandemly (one after another); such a series of repeats in the genome is called a **short tandem repeat (STR)**. For example, one person might have the sequence AGAT repeated 12 times in a row at one place in the genome, the sequence GATA repeated 35 times at a second place, and so on; another person is likely to have the same sequences at the same places but with a different number of repeats. By focusing on STRs, forensic scientists are able to compare the tiny fraction of the genome that is most likely to vary from person to person.

STR analysis is a method of DNA profiling that compares the lengths of STR sequences at specific sites in the genome. The standard STR analysis procedure used by U.S. law enforcement agencies compares the number of repeats of specific four-nucleotide DNA sequences at 13 sites scattered throughout the genome. Each repeat site, which typically contains from 3 to 50 four-nucleotide repeats in a row, varies widely from person to person. In fact, some STRs used in the standard procedure have up to 80 variations in the number of repeats. In the United States, the number of repeats at each site is entered into a database called CODIS (Combined DNA Index System) administered by the Federal Bureau of Investigation. Law enforcement agencies around the world can access CODIS to search for matches to DNA samples they have obtained from crime scenes or suspects.

Consider the two samples of DNA shown in **Figure 12.14**. Imagine that the top DNA segment was obtained at a crime scene and the bottom from a suspect's blood. The two segments have the same number of repeats at the first site: 7 repeats of the four-nucleotide DNA sequence AGAT (in orange). However, there isn't a match at the second site: 8 repeats of GATA (in purple) in the crime scene DNA, compared with 12 repeats in the suspect's DNA. To create a DNA profile, a scientist uses PCR to specifically amplify the regions of DNA that include these STR sites. The resulting fragments are then compared. In this case, that comparison reveals that the two samples of DNA could not have come from the same person. ☑

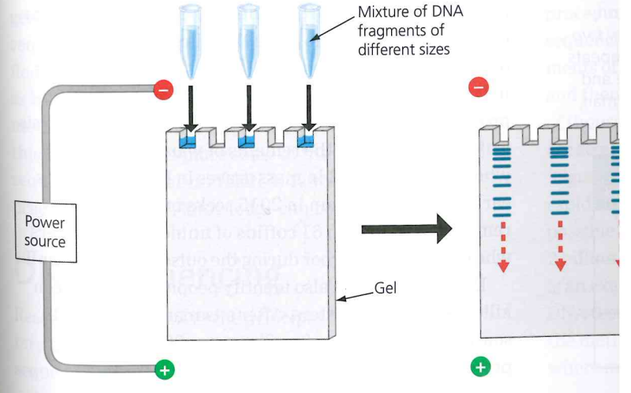


▶ **Figure 12.14 Short tandem repeat (STR) sites.** Scattered throughout the genome, STR sites contain tandem repeats of four-nucleotide sequences. The number of repetitions at each site can vary from individual to individual. In this figure, both DNA samples have the same number of repeats (7) at the first STR site, but different numbers (8 versus 12) at the second.

Gel Electrophoresis

DNA profiling by STR analysis depends upon comparing lengths of DNA fragments. This can be accomplished by **gel electrophoresis**, a method for sorting macromolecules—usually proteins or nucleic acids—primarily by electrical charge and size. **Figure 12.15** shows how gel electrophoresis separates DNA fragments obtained from different sources. A sample with many copies of the DNA from each source is placed in a separate well (hole) at one end of a gel, a thin slab of jellylike material that acts as a molecular sieve. A negatively charged electrode is then attached to this end of the gel and a positive electrode to the other end. Because the phosphate (PO_4^-) groups of

▼ **Figure 12.15 Gel electrophoresis of DNA molecules.** The photo shows DNA fragments of various sizes visibly stained on the gel.



exponentially growing population of identical DNA
ecules. The key to automated PCR is an unusually
stable DNA polymerase, first isolated from pro-
otes living in hot springs (see Figure 15.16). Unlike
proteins—which denature, or fall apart, at high
peratures—this enzyme can withstand the heat at the
of each cycle.

NA molecules are typically very long, but usually
a very small target region of a large DNA mol-
e needs to be amplified. The key to amplifying one
ticular segment of DNA and no others is the use
rimers, short (usually 15–20 nucleotides long),
mically synthesized single-stranded DNA mol-
les. For each experiment, specific primers are cho-
that are complementary to sequences found only at
h end of the target sequence. The primers thus bind
sequences that flank the target sequence, marking
start and end points for the segment of DNA to be
plified. Beginning with a single DNA molecule and
e appropriate primers, automated PCR can generate
ndreds of billions of copies of the desired sequence
a few hours.

In addition to forensic applications, PCR can be used
the treatment and diagnosis of disease. For example,
ecause the sequence of the genome of HIV (the virus
iat causes AIDS) is known, PCR can be used to amplify,
nd thus detect, HIV in blood or tissue samples. In fact,
CR is often the best way to detect this otherwise elusive
irus. Medical scientists can now diagnose hundreds of
uman genetic disorders by using PCR with primers that
arget the genes associated with these disorders. The
mplified DNA product is then studied to reveal the pres-
nce or absence of the disease-causing mutation. Among
he genes for human diseases that have been identified
are those for sickle-cell disease, hemophilia, cystic fibro-
sis, Huntington's disease, and Duchenne muscular dys-
trophy. Individuals afflicted with such diseases can often
be identified before the onset of symptoms, even before
birth, allowing for preventative medical care to begin.
PCR can also be used to identify symptomless carriers of
potentially harmful recessive alleles (see Figure 9.14).
Parents may thus be informed of whether they have a
risk of bearing a child with a rare disease that they do not
themselves display.

Short Tandem Repeat (STR) Analysis

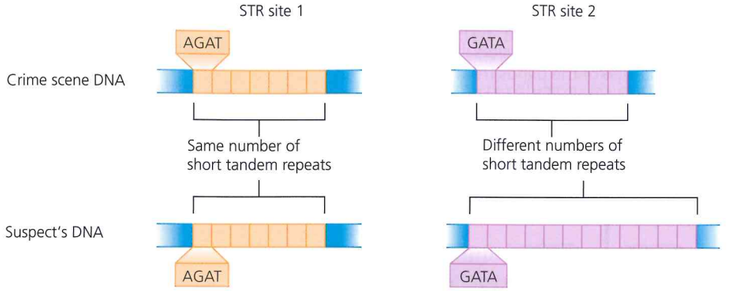
How do you prove that two samples of DNA come from the same person? You could compare the entire genomes found in the two samples. But such an approach is impractical because the DNA of two humans of the same sex is 99.9% identical. Instead, forensic scientists typically compare about a dozen short segments of non-coding repetitive DNA that are known to vary from person to person. Have you ever seen a puzzle in a magazine that presents two nearly identical photos and asks you to find the few differences between them? In a similar way, scientists can focus on the few areas of difference in the human genome, ignoring the identical majority.

Repetitive DNA, which makes up much of the DNA that lies between genes in humans, consists of nucleotide sequences that are present in multiple copies in the genome. Some of this DNA consists of short sequences repeated many times tandemly (one after another); such a series of repeats in the genome is called a **short tandem repeat (STR)**. For example, one person might have the sequence AGAT repeated 12 times in a row at one place in the genome, the sequence GATA repeated 35 times at a second place, and so on; another person is likely to have the same sequences at the same places but with a different number of repeats. By focusing on STRs, forensic scientists are able to compare the tiny fraction of the genome that is most likely to vary from person to person.

STR analysis is a method of DNA profiling that compares the lengths of STR sequences at specific sites in the genome. The standard STR analysis procedure used by U.S. law enforcement agencies compares the number of repeats of specific four-nucleotide DNA sequences at 13 sites scattered throughout the genome. Each repeat site, which typically contains from 3 to 50 four-nucleotide repeats in a row, varies widely from person to person. In fact, some STRs used in the standard procedure have up to 80 variations in the number of repeats. In the United States, the number of repeats at each site is entered into a database called CODIS (Combined DNA Index System) administered by the Federal Bureau of Investigation. Law enforcement agencies around the world can access CODIS to search for matches to DNA samples they have obtained from crime scenes or suspects.

Consider the two samples of DNA shown in Figure 12.14. Imagine that the top DNA segment was obtained at a crime scene and the bottom from a suspect's blood. The two segments have the same number of repeats at the first site: 7 repeats of the four-nucleotide DNA sequence AGAT (in orange). However, there isn't a match at the second site: 8 repeats of GATA (in purple) in the crime scene DNA, compared with 12 repeats in the suspect's DNA. To create a DNA profile, a scientist uses PCR to specifically amplify the regions of DNA that include these STR sites. The resulting fragments are then compared. In this case, that comparison reveals that the two samples of DNA could not have come from the same person. ✅

▶ **Figure 12.14 Short tandem repeat (STR) sites.** Scattered throughout the genome, STR sites contain tandem repeats of four-nucleotide sequences. The number of repetitions at each site can vary from individual to individual. In this figure, both DNA samples have the same number of repeats (7) at the first STR site, but different numbers (8 versus 12) at the second.



Gel Electrophoresis

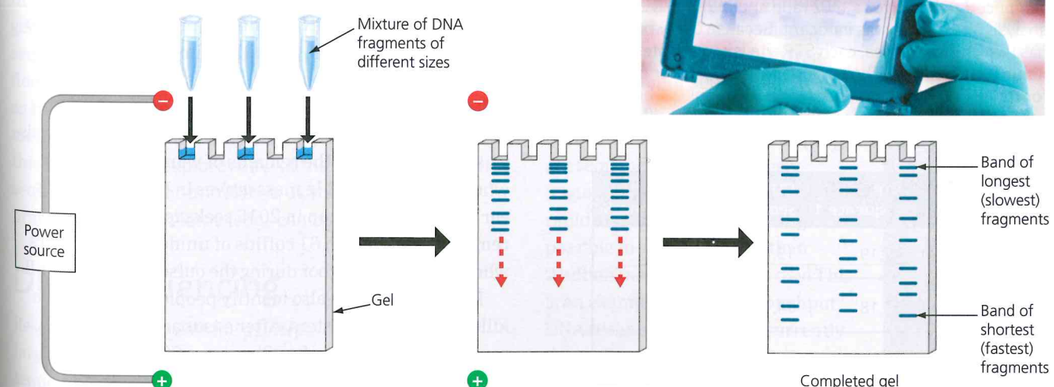
DNA profiling by STR analysis depends upon comparing lengths of DNA fragments. This can be accomplished by **gel electrophoresis**, a method for sorting macromolecules—usually proteins or nucleic acids—primarily by electrical charge and size. Figure 12.15 shows how gel electrophoresis separates DNA fragments obtained from different sources. A sample with many copies of the DNA from each source is placed in a separate well (hole) at one end of a gel, a thin slab of jellylike material that acts as a molecular sieve. A negatively charged electrode is then attached to this end of the gel and a positive electrode to the other end. Because the phosphate (PO_4^-) groups of nucleotides give DNA fragments a negative charge, the fragments move through the gel toward the positive pole. Imagine a small animal scampering through a thicket of vines, while a large animal plods much more slowly. Similarly, shorter DNA fragments move farther through a gel than do longer DNA fragments. Gel electrophoresis thus separates DNA fragments by length. When the current is turned off, a series of bands (blue smudges in the photo)is left in each column of the gel. Each band is a collection of DNA fragments of the same length. The DNA bands can be made visible by staining, by exposure onto photographic film (if radioactively labeled), or by measuring fluorescence (if labeled with a fluorescent dye). ✅

✅ CHECKPOINT

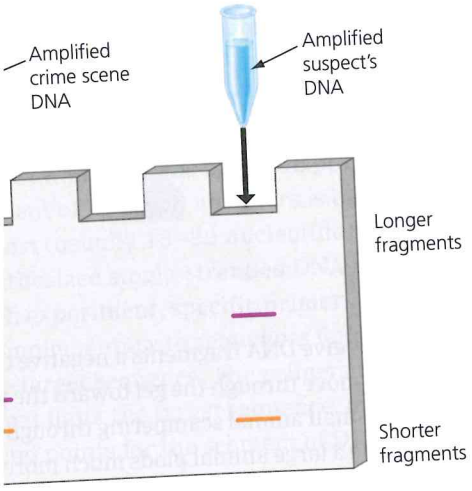
You use a restriction enzyme to cut a long DNA molecule that has three copies of the enzyme's recognition sequence clustered near one end. When you separate the restriction fragments by gel electrophoresis, how do you expect the bands to appear?

■ *Answer: three bands (short fragments) near the positive pole and one band (long fragment) near the negative pole*

▼ **Figure 12.15 Gel electrophoresis of DNA molecules.** The photo shows DNA fragments of various sizes visibly stained on the gel.



sualizing STR fragment patterns.
the bands that would result from gel
the STR sites illustrated in Figure 12.14.
the bands from the crime scene DNA
ne of the bands from the suspect's DNA.



shows the gel that would result from using
esis to separate the DNA fragments from the
ure 12.14. This figure simplifies the process;
analysis uses more than two sites and uses
thod to visualize the results. The differences
s of the bands reflect the different lengths of
nents. This gel would provide evidence that
ne DNA did not come from the suspect.
ing can provide evidence of guilt or inno-
017, lawyers at the Innocence Project, a non-
ganization in New York City, have helped to
re than 350 convicted criminals, including
on death row. The average sentence served by
ere exonerated was 14 years. In nearly half of
NA profiling has also identified the true per-
gure 12.17 presents some data from a real case

17 DNA profiling: proof of innocence and guilt.
Washington (shown at left) was convicted and
death for a 1982 rape and murder. In 2000, STR
ved he was innocent. Because every person
omosomes, each STR site is represented by two
epeats. The table shows the number of repeats
R DNA sequences in three samples. These and
ata exonerated Washington and led other man,
sley, to plead guilty to the murder.

of	STR sequence 1	STR sequence 2	STR sequence3
victim	17,19	13,16	12,12
ington	16,18	14,15	11,12
insley	17,19	13,16	12,12

in which STR analysis proved a convicted man innocent and helped identify the true perpetrator.

Just how reliable is a genetic profile? In forensic cases using STR analysis with the 13 standard sites, the probability of two people having identical DNA profiles is somewhere between one chance in 10 billion and one in several trillion. (The exact probability depends on the frequency of the particular DNA sequences in the general population.) Thus, despite problems that can still arise from insufficient data, human error, or flawed evidence, genetic profiles are now accepted as compelling evidence by legal experts and scientists alike.

Investigating Murder, Paternity, and Ancient DNA

Since DNA profiling was introduced in 1986, it has become a standard tool of forensics and has provided crucial evidence in many famous investigations. After the death of terrorist leader Osama bin Laden in 2011, U.S. Special Forces members obtained a sample of his DNA. Within hours, a military laboratory in Afghanistan had compared the tissue against samples obtained from several of bin Laden's relatives, including a sister who had died of brain cancer in a Boston hospital in 2010. Although facial recognition and an eyewitness identification provided preliminary evidence, it was DNA that provided a conclusive match, officially ending the long hunt for the notorious terrorist.

DNA profiling can also be used to identify murder victims. The largest such effort took place after the World Trade Center attack on September 11, 2001. Forensic scientists in New York City worked for years to identify more than 20,000 samples of victims' remains. DNA profiles of tissue samples from the disaster site were matched to DNA profiles from tissue known to be from the victims or their relatives. More than half of the victims identified at the World Trade Center site were recognized solely by DNA evidence, providing closure to many grieving families. Since that time, the victims of other atrocities, such as mass killings during civil wars in Europe and Africa, have been identified using DNA profiling techniques. In 2010, for example, DNA analysis was used to identify the remains of war crime victims who had been buried in mass graves in Bosnia 15 years earlier. An effort begun in 2015 seeks to identify the remains contained in 61 coffins of unidentified soldiers who died at Pearl Harbor during the outset of World War II.

DNA profiling can also identify people who have been killed in natural disasters. After a tsunami devastated southern Asia the day after Christmas 2004, DNA profiling identified hundreds of victims.

Comparing the DNA of a mo
purported father can settle a q
Sometimes paternity is of hist
profiling proved that Thomas
relative fathered a child with a
Hemings. Similarly, tests conc
a woman's story that Warren
her child while he was preside
case, researchers investigated
of Marie Antoinette **(Figure 12.**
France, survived the French R

▼ Figure 12.18 Marie Antoinette
Louis (depicted with his mother ir
of the Queen of France, did not s

Bioinforma

In the past decade, new experi
generated enormous volumes
sequences. The need to make s
flood of information has given
as **bioinformatics**, the applica
methods to the storage and ana
this section, we'll explore some
sequence data are accumulate
practical ways such knowledge

DNA Sequencing

Researchers can exploit the pri
tary base pairing to determine
sequence of small DNA molecu

▼ Figure 12.16 Visualizing STR fragment patterns. This figure shows the bands that would result from gel electrophoresis of the STR sites illustrated in Figure 12.14. Notice that one of the bands from the crime scene DNA does not match one of the bands from the suspect's DNA.

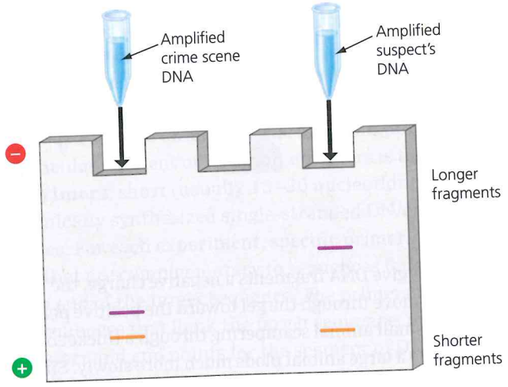
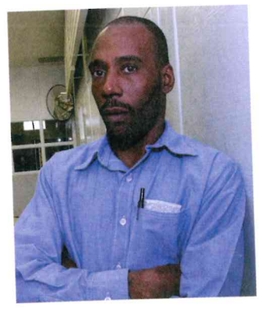


Figure 12.16 shows the gel that would result from using gel electrophoresis to separate the DNA fragments from the example in Figure 12.14. This figure simplifies the process; an actual STR analysis uses more than two sites and uses a different method to visualize the results. The differences in the locations of the bands reflect the different lengths of the DNA fragments. This gel would provide evidence that the crime scene DNA did not come from the suspect.

DNA profiling can provide evidence of guilt or innocence. As of 2017, lawyers at the Innocence Project, a nonprofit legal organization in New York City, have helped to exonerate more than 350 convicted criminals, including 20 who were on death row. The average sentence served by those who were exonerated was 14 years. In nearly half of these cases, DNA profiling has also identified the true perpetrators. **Figure 12.17** presents some data from a real case in which STR analysis proved a convicted man innocent and helped identify the true perpetrator.



▼ Figure 12.17 DNA profiling: proof of innocence and guilt. In 1984, Earl Washington (shown at left) was convicted and sentenced to death for a 1982 rape and murder. In 2000, STR analysis showed he was innocent. Because every person has two chromosomes, each STR site is represented by two numbers of repeats. The table shows the number of repeats for three STR DNA sequences in three samples. These and other STR data exonerated Washington and led other man, Kenneth Tinsley, to plead guilty to the murder.

Source of sample	STR sequence 1	STR sequence 2	STR sequence3
Semen on victim	17,19	13,16	12,12
Earl Washington	16,18	14,15	11,12
Kenneth Tinsley	17,19	13,16	12,12

Just how reliable is a genetic profile? In forensic cases using STR analysis with the 13 standard sites, the probability of two people having identical DNA profiles is somewhere between one chance in 10 billion and one in several trillion. (The exact probability depends on the frequency of the particular DNA sequences in the general population.) Thus, despite problems that can still arise from insufficient data, human error, or flawed evidence, genetic profiles are now accepted as compelling evidence by legal experts and scientists alike.

Investigating Murder, Paternity, and Ancient DNA

Since DNA profiling was introduced in 1986, it has become a standard tool of forensics and has provided crucial evidence in many famous investigations. After the death of terrorist leader Osama bin Laden in 2011, U.S. Special Forces members obtained a sample of his DNA. Within hours, a military laboratory in Afghanistan had compared the tissue against samples obtained from several of bin Laden's relatives, including a sister who had died of brain cancer in a Boston hospital in 2010. Although facial recognition and an eyewitness identification provided preliminary evidence, it was DNA that provided a conclusive match, officially ending the long hunt for the notorious terrorist.

DNA profiling can also be used to identify murder victims. The largest such effort took place after the World Trade Center attack on September 11, 2001. Forensic scientists in New York City worked for years to identify more than 20,000 samples of victims' remains. DNA profiles of tissue samples from the disaster site were matched to DNA profiles from tissue known to be from the victims or their relatives. More than half of the victims identified at the World Trade Center site were recognized solely by DNA evidence, providing closure to many grieving families. Since that time, the victims of other atrocities, such as mass killings during civil wars in Europe and Africa, have been identified using DNA profiling techniques. In 2010, for example, DNA analysis was used to identify the remains of war crime victims who had been buried in mass graves in Bosnia 15 years earlier. An effort begun in 2015 seeks to identify the remains contained in 61 coffins of unidentified soldiers who died at Pearl Harbor during the outset of World War II.

DNA profiling can also identify people who have been killed in natural disasters. After a tsunami devastated southern Asia the day after Christmas 2004, DNA profiling identified hundreds of victims.

Comparing the DNA of a mother, her child, and a purported father can settle a question of paternity. Sometimes paternity is of historical interest: DNA profiling proved that Thomas Jefferson or a close male relative fathered a child with an enslaved woman, Sally Hemings. Similarly, tests conducted in 2015 confirmed a woman's story that Warren G. Harding had fathered her child while he was president. In another historical case, researchers investigated whether any descendants of Marie Antoinette **(Figure 12.18)**, one-time queen of France, survived the French Revolution. DNA extracted from a preserved heart said to belong to h compared with DNA extracted from a lock hair. A DNA match proved that her last kn in fact, died in jail during the revolution. a former backup singer for James Brown, of Soul," sued his estate after his death, cl her child was Brown's son. A DNA paterni her claim, and 25% of Brown's estate was the mother and child.

▼ Figure 12.18 Marie Antoinette. DNA profiling proved that Louis (depicted with his mother in this 1785 painting), the son of the Queen of France, did not survive the French Revolution.

DNA profiling can also help protect endan by proving the origin of contraband animal example, analysis of seized elephant tusks c location of the poaching, allowing enforcem increase surveillance and prosecute those re 2014, three tiger poachers in India were sen years in jail after DNA profiling matched the flesh to tissue under the poachers' fingernai

Modern methods of DNA profiling are so powerful that the DNA samples can be in a p degraded state. Such advances are revolutio study of ancient remains. For example, a 20 DNA extracted from five mummified Egypti (dating from 800 B.C. to 100 A.D.) was used t geographic origins of the mummies, as well a pathogens that had infected those people be Another study determined that DNA extract 27,000-year-old Siberian mammoth was 98. to DNA from living African elephants. Other ing a large collection of mammoth samples s the last populations of the huge beasts migra North America to Siberia, where separate sp before dying out several thousand years ago.

Bioinformatics

In the past decade, new experimental techniques have generated enormous volumes of data related to DNA sequences. The need to make sense of an ever-increasing flood of information has given rise to the field known as **bioinformatics**, the application of computational methods to the storage and analysis of biological data. In this section, we'll explore some of the methods by which sequence data are accumulated, as well as many of the practical ways such knowledge can be used.

DNA Sequencing

Researchers can exploit the principle of complementary base pairing to determine the complete nucleotide sequence of small DNA molecules. This process is known as **DNA sequencing**. In one standard procedure, called "next-generation sequencing," DNA is cut into fragments of around 300 nucleotides, and then thousands or hundreds of thousands of these fragments are sequenced simultaneously **(Figure 12.19)**. This technology is rapid and inexpensive, making it possible to sequence more than 2 billion nucleotides in one day! This is an example of "high-throughput" DNA technology, which is currently the method of choice for studies where massive numbers of DNA

igure 12.16 Visualizing STR fragment patterns.
s figure shows the bands that would result from gel
:trophoresis of the STR sites illustrated in Figure 12.14.
tice that one of the bands from the crime scene DNA
es not match one of the bands from the suspect's DNA.

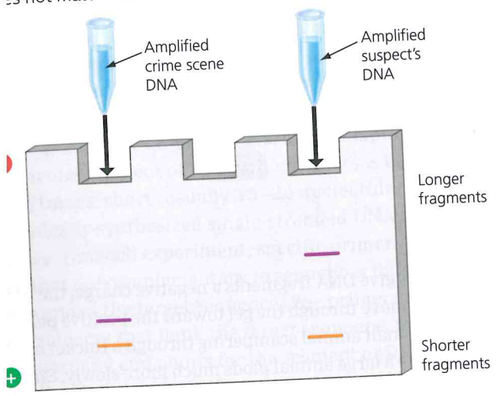


Figure 12.16 shows the gel that would result from using
gel electrophoresis to separate the DNA fragments from the
example in Figure 12.14. This figure simplifies the process;
an actual STR analysis uses more than two sites and uses
a different method to visualize the results. The differences
in the locations of the bands reflect the different lengths of
the DNA fragments. This gel would provide evidence that
the crime scene DNA did not come from the suspect.

DNA profiling can provide evidence of guilt or inno-
cence. As of 2017, lawyers at the Innocence Project, a non-
profit legal organization in New York City, have helped to
exonerate more than 350 convicted criminals, including
20 who were on death row. The average sentence served by
those who were exonerated was 14 years. In nearly half of
these cases, DNA profiling has also identified the true per-
petrators. **Figure 12.17** presents some data from a real case

▼ Figure 12.17 DNA profiling: proof of innocence and guilt.
In 1984, Earl Washington (shown at left) was convicted and sentenced to death for a 1982 rape and murder. In 2000, STR analysis showed he was innocent. Because every person has two chromosomes, each STR site is represented by two numbers of repeats. The table shows the number of repeats for three STR DNA sequences in three samples. These and other STR data exonerated Washington and led other man, Kenneth Tinsley, to plead guilty to the murder.

Source of sample	STR sequence 1	STR sequence 2	STR sequence3
Semen on victim	17,19	13,16	12,12
Earl Washington	16,18	14,15	11,12
Kenneth Tinsley	17,19	13,16	12,12

in which STR analysis proved a convicted man innocent and helped identify the true perpetrator.

Just how reliable is a genetic profile? In forensic cases using STR analysis with the 13 standard sites, the probability of two people having identical DNA profiles is somewhere between one chance in 10 billion and one in several trillion. (The exact probability depends on the frequency of the particular DNA sequences in the general population.) Thus, despite problems that can still arise from insufficient data, human error, or flawed evidence, genetic profiles are now accepted as compelling evidence by legal experts and scientists alike.

Investigating Murder, Paternity, and Ancient DNA

Since DNA profiling was introduced in 1986, it has become a standard tool of forensics and has provided crucial evidence in many famous investigations. After the death of terrorist leader Osama bin Laden in 2011, U.S. Special Forces members obtained a sample of his DNA. Within hours, a military laboratory in Afghanistan had compared the tissue against samples obtained from several of bin Laden's relatives, including a sister who had died of brain cancer in a Boston hospital in 2010. Although facial recognition and an eyewitness identification provided preliminary evidence, it was DNA that provided a conclusive match, officially ending the long hunt for the notorious terrorist.

DNA profiling can also be used to identify murder victims. The largest such effort took place after the World Trade Center attack on September 11, 2001. Forensic scientists in New York City worked for years to identify more than 20,000 samples of victims' remains. DNA profiles of tissue samples from the disaster site were matched to DNA profiles from tissue known to be from the victims or their relatives. More than half of the victims identified at the World Trade Center site were recognized solely by DNA evidence, providing closure to many grieving families. Since that time, the victims of other atrocities, such as mass killings during civil wars in Europe and Africa, have been identified using DNA profiling techniques. In 2010, for example, DNA analysis was used to identify the remains of war crime victims who had been buried in mass graves in Bosnia 15 years earlier. An effort begun in 2015 seeks to identify the remains contained in 61 coffins of unidentified soldiers who died at Pearl Harbor during the outset of World War II.

DNA profiling can also identify people who have been killed in natural disasters. After a tsunami devastated southern Asia the day after Christmas 2004, DNA profiling identified hundreds of victims.

Comparing the DNA of a mother, her child, and a purported father can settle a question of paternity. Sometimes paternity is of historical interest: DNA profiling proved that Thomas Jefferson or a close male relative fathered a child with an enslaved woman, Sally Hemings. Similarly, tests conducted in 2015 confirmed a woman's story that Warren G. Harding had fathered her child while he was president. In another historical case, researchers investigated whether any descendants of Marie Antoinette **(Figure 12.18)**, one-time queen of France, survived the French Revolution. DNA extracted from a preserved heart said to belong to her son was compared with DNA extracted from a lock of Marie's hair. A DNA match proved that her last known heir had, in fact, died in jail during the revolution. More recently, a former backup singer for James Brown, the "Godfather of Soul," sued his estate after his death, claiming that her child was Brown's son. A DNA paternity test proved her claim, and 25% of Brown's estate was awarded to the mother and child.

▼ Figure 12.18 Marie Antoinette. DNA profiling proved that Louis (depicted with his mother in this 1785 painting), the son of the Queen of France, did not survive the French Revolution.

DNA profiling can also help protect endangered species by proving the origin of contraband animal products. For example, analysis of seized elephant tusks can pinpoint the location of the poaching, allowing enforcement officials to increase surveillance and prosecute those responsible. In 2014, three tiger poachers in India were sentenced to five years in jail after DNA profiling matched the dead tigers' flesh to tissue under the poachers' fingernails.

Modern methods of DNA profiling are so specific and powerful that the DNA samples can be in a partially degraded state. Such advances are revolutionizing the study of ancient remains. For example, a 2014 study of DNA extracted from five mummified Egyptian heads (dating from 800 B.C. to 100 A.D.) was used to deduce the geographic origins of the mummies, as well as to identify pathogens that had infected those people before they died. Another study determined that DNA extracted from a 27,000-year-old Siberian mammoth was 98.6% identical to DNA from living African elephants. Other studies involving a large collection of mammoth samples suggested that the last populations of the huge beasts migrated from North America to Siberia, where separate species interbred before dying out several thousand years ago.

Bioinformatics

In the past decade, new experimental techniques have generated enormous volumes of data related to DNA sequences. The need to make sense of an ever-increasing flood of information has given rise to the field known as **bioinformatics**, the application of computational methods to the storage and analysis of biological data. In this section, we'll explore some of the methods by which sequence data are accumulated, as well as many of the practical ways such knowledge can be used.

DNA Sequencing

Researchers can exploit the principle of complementary base pairing to determine the complete nucleotide sequence of small DNA molecules. This process is known as **DNA sequencing**. In one standard procedure, called "next-generation sequencing," DNA is cut into fragments of around 300 nucleotides, and then thousands or hundreds of thousands of these fragments are sequenced simultaneously **(Figure 12.19)**. This technology is rapid and inexpensive, making it possible to sequence more than 2 billion nucleotides in one day! This is an example of "high-throughput" DNA technology, which is currently the method of choice for studies where massive numbers of DNA

▼ Figure 12.19 A DNA sequencer.
This high-throughput DNA-sequencing machine can process half a billion bases in a single 10-hour run.

samples—even representing an entire genome or a collection of genomes—are being sequenced.

More recently, scientists have improved or replaced next-generation sequencing. Several groups have been working on "third-generation sequencing," wherein a single, very long DNA molecule is sequenced on its own. The idea is to move a single strand of DNA through a very small pore in a membrane (a nanopore) while administering an electrical current that will detect the nitrogenous bases one by one. For each type of base, the electrical current is interrupted for a slightly different length of time, allowing the base sequence to be determined nucleotide by nucleotide.

In 2015, the first nanopore sequencer went on the market; this device is the size of a small candy bar and connects to a computer through a USB port. Software allows the immediate identification and analysis of the sequence. This is one of many approaches to increase the rate and cut the cost of sequencing, while allowing the methodology to move out of the laboratory and into the field.

Next-generation and third-generation sequencing techniques are ushering in a new era of faster, more affordable sequencing. Taken to their logical extreme, these techniques can be applied to whole genomes—our next topic.

✓ CHECKPOINT

Based on Figure 12.20, about how many nucleotides and genes are contained in the human genome?

■ Answer: about 3 billion nucleotides and 21,000 genes

Genomics

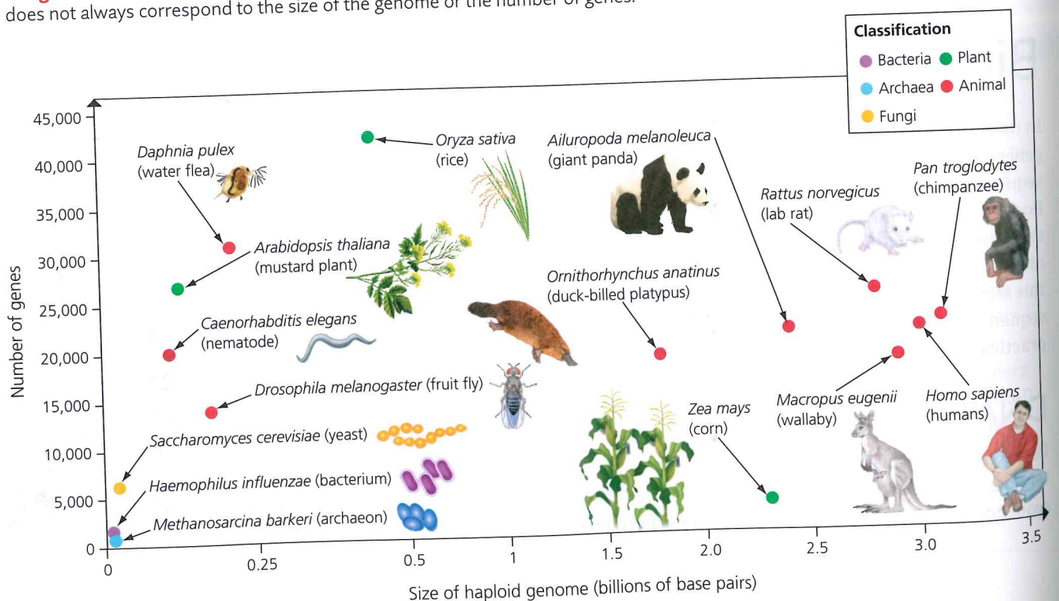
Improved DNA-sequencing techniques have transformed how we explore fundamental biological questions about evolution and how life works. A major leap forward occurred in 1995 when a team of scientists determined the nucleotide sequence of the entire genome of *Haemophilus influenzae*, a bacterium that causes several human diseases, including pneumonia and meningitis. **Genomics**, the study of complete sets of genes (genomes), was born.

The first targets of genomics research were bacteria, which have relatively little DNA. Researchers then studied more complex organisms with much larger genomes. Baker's yeast (*Saccharomyces cerevisiae*) was the first eukaryote to have its full sequence determined, and the roundworm *Caenorhabditis elegans* was the first multicellular organism. Other sequenced animals include the fruit fly (*Drosophila melanogaster*) and lab rat (*Rattus norvegicus*), both model organisms for genetics research. Among the sequenced plants are *Arabidopsis thaliana*, a type of mustard plant used as a model organism, and rice (*Oryza sativa*), one of the most economically important crops.

The genomes of thousands of species have been published, and tens of thousands more are in progress **(Figure 12.20)**. The majority of organisms sequenced are prokaryotes, including more than 4,000 bacterial species and nearly 200 archaea. Hundreds of eukaryotic genomes—including protists, fungi, plants, and animals both invertebrate and vertebrate—have been completed. Genome sequences have been determined for cells from several cancers, for ancient humans, and for the many bacteria that live in the human intestine. **As the ultimate repository of the genetic information from which all of life's inherited characteristics develop, genomes hold the key to our genetic identity.** ✓

VISUALIZING THE DATA

▼ Figure 12.20 Comparing whole genomes. Notice that the size or complexity of an organism does not always correspond to the size of the genome or the number of genes.



Genome-Mapping Techniques

Genomes are most often sequenced using a technique called the **whole-genome shotgun method (Figure 12.21)**. The first step is to chop the entire genome into fragments using restriction enzymes. Next, all the fragments are cloned and sequenced. Finally, computers running specialized mapping software reassemble the millions of overlapping short sequences into a single continuous sequence for every chromosome—an entire genome.

The DNA sequences determined by many research groups in the United States are deposited in GenBank, a database available to anyone through the Internet. You can browse it yourself at the National Center for Biotechnology Information: www.ncbi.nlm.nih.gov. As of 2017, GenBank includes the sequences of more than 200 billion base pairs of DNA! The database is constantly updated, and the amount of data it contains doubles every 18 months. Any sequence in the database can be retrieved and analyzed. For example, software can compare a collection of sequences from different species and diagram them as an evolutionary tree based on the sequence relationships. Bioinformatics has thereby revolutionized evolutionary biology by opening a vast new reservoir of data that can test evolutionary hypotheses. Next, we'll discuss a particularly notable example of a sequenced animal genome—our own.

The Human Genome

The **Human Genome Project** was a massive scientific endeavor to determine the nucleotide sequence of all the DNA in the human genome and to identify the location and sequence of every gene. The project began in 1990 as an effort by government-funded researchers from six nations. Several years into the project, private companies joined the effort. At the completion of the project, more than 99% of the genome had been determined to 99.999% accuracy. (There remain a few hundred gaps of unknown sequence

that require special
tious project has pr
minate the genetic b

The chromosome
plus the X and Y sex
3 billion nucleotide
sequence printed in
see on this page, the
stories high! Howev
Genome Project is t
genes—currently est
to the number found

Like the genomes
small amount of tota
code for proteins, tR
otes have a huge am

▼ Figure 12.21 Genc
bottom, a technician
shotgun method (dep

samples—even representing an entire genome or a collection of genomes—are being sequenced.

More recently, scientists have improved or replaced next-generation sequencing. Several groups have been working on "third-generation sequencing," wherein a single, very long DNA molecule is sequenced on its own. The idea is to move a single strand of DNA through a very small pore in a membrane (a nanopore) while administering an electrical current that will detect the nitrogenous bases one by one. For each type of base, the electrical current is interrupted for a slightly different length of time, allowing the base sequence to be determined nucleotide by nucleotide.

In 2015, the first nanopore sequencer went on the market; this device is the size of a small candy bar and connects to a computer through a USB port. Software allows the immediate identification and analysis of the sequence. This is one of many approaches to increase the rate and cut the cost of sequencing, while allowing the methodology to move out of the laboratory and into the field.

Next-generation and third-generation sequencing techniques are ushering in a new era of faster, more affordable sequencing. Taken to their logical extreme, these techniques can be applied to whole genomes—our next topic.

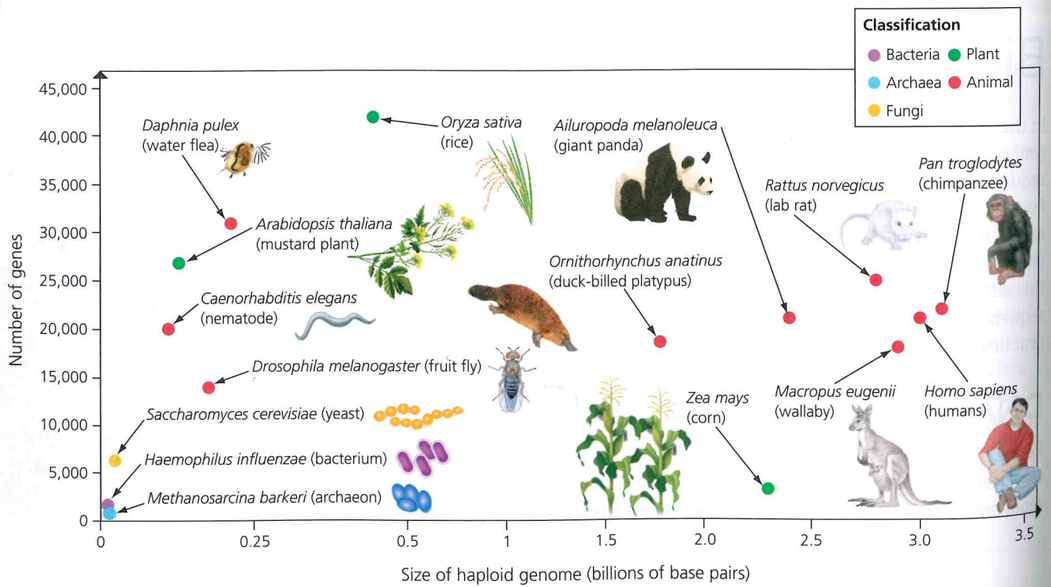
Genomics

Improved DNA-sequencing techniques have transformed how we explore fundamental biological questions about evolution and how life works. A major leap forward occurred in 1995 when a team of scientists determined the nucleotide sequence of the entire genome of *Haemophilus influenzae*, a bacterium that causes several human diseases, including pneumonia and meningitis. **Genomics**, the study of complete sets of genes (genomes), was born.

The first targets of genomics research were bacteria, which have relatively little DNA. Researchers then studied more complex organisms with much larger genomes. Baker's yeast (*Saccharomyces cerevisiae*) was the first eukaryote to have its full sequence determined, and the roundworm *Caenorhabditis elegans* was the first multicellular organism. Other sequenced animals include the fruit fly (*Drosophila melanogaster*) and lab rat (*Rattus norvegicus*), both model organisms for genetics research. Among the sequenced plants are *Arabidopsis thaliana*, a type of mustard plant used as a model organism, and rice (*Oryza sativa*), one of the most economically important crops.

The genomes of thousands of species have been published, and tens of thousands more are in progress (Figure 12.20). The majority of organisms sequenced are prokaryotes, including more than 4,000 bacterial species and nearly 200 archaea. Hundreds of eukaryotic genomes—including protists, fungi, plants, and animals both invertebrate and vertebrate—have been completed. Genome sequences have been determined for cells from several cancers, for ancient humans, and for the many bacteria that live in the human intestine. As the ultimate repository of the genetic information from which all of life's inherited characteristics develop, genomes hold the key to our genetic identity. ✅

▼ **Figure 12.20 Comparing whole genomes.** Notice that the size or complexity of an organism does not always correspond to the size of the genome or the number of genes.



Genome-Mapping Techniques

Genomes are most often sequenced using a technique called the **whole-genome shotgun method** (Figure 12.21). The first step is to chop the entire genome into fragments using restriction enzymes. Next, all the fragments are cloned and sequenced. Finally, computers running specialized mapping software reassemble the millions of overlapping short sequences into a single continuous sequence for every chromosome—an entire genome.

The DNA sequences determined by many research groups in the United States are deposited in GenBank, a database available to anyone through the Internet. You can browse it yourself at the National Center for Biotechnology Information: www.ncbi.nlm.nih.gov. As of 2017, GenBank includes the sequences of more than 200 billion base pairs of DNA! The database is constantly updated, and the amount of data it contains doubles every 18 months. Any sequence in the database can be retrieved and analyzed. For example, software can compare a collection of sequences from different species and diagram them as an evolutionary tree based on the sequence relationships. Bioinformatics has thereby revolutionized evolutionary biology by opening a vast new reservoir of data that can test evolutionary hypotheses. Next, we'll discuss a particularly notable example of a sequenced animal genome—our own.

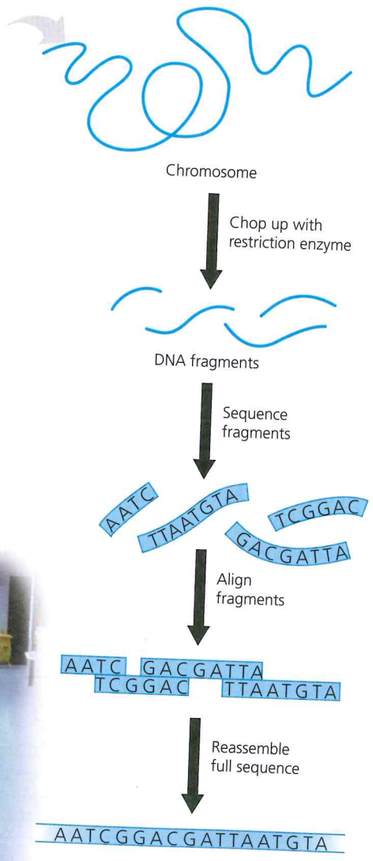
The Human Genome

The **Human Genome Project** was a massive scientific endeavor to determine the nucleotide sequence of all the DNA in the human genome and to identify the location and sequence of every gene. The project began in 1990 as an effort by government-funded researchers from six nations. Several years into the project, private companies joined the effort. At the completion of the project, more than 99% of the genome had been determined to 99.999% accuracy. (There remain a few hundred gaps of unknown sequence that require special methods to figure out.) This ambitious project has provided a wealth of data that may illuminate the genetic basis of what it means to be human.

The chromosomes in the human genome (22 autosomes plus the X and Y sex chromosomes) contain approximately 3 billion nucleotide pairs of DNA. If you imagine this sequence printed in letters (A, T, C, and G) the same size you see on this page, the sequence would fill a stack of books 18 stories high! However, the biggest surprise from the Human Genome Project is the relatively small number of human genes—currently estimated to be about 21,000—very close to the number found in a roundworm!

Like the genomes of most complex eukaryotes, only a small amount of total human DNA consists of genes that code for proteins, tRNAs, or rRNAs. Most complex eukaryotes have a huge amount of noncoding DNA—about 98.5%

▼ **Figure 12.21 Genome sequencing.** In the photo at the bottom, a technician performs a step in the whole-genome shotgun method (depicted in the diagram).



YOU HAVE ABOUT THE SAME NUMBER OF GENES AS A WORM, AND ONLY HALF AS MANY AS A RICE PLANT.

of human DNA is of this type. Some of this noncoding DNA is made up of gene control sequences such as promoters, enhancers, and microRNAs (see Chapter 11). Other noncoding regions include introns and repetitive DNA (some of which is used in DNA profiling). Some noncoding DNA is important to our health, with certain regions known to carry disease-causing mutations. But the function of most noncoding DNA remains unknown.

The human genome sequenced by government-funded scientists was actually a reference genome compiled from a group of individuals. As of today, the complete genomes of many individuals have been completed. Whereas sequencing the first human genome took 13 years and cost $100 million, we are rapidly approaching the day when an individual's genome can be sequenced in a matter of hours for less than $1,000.

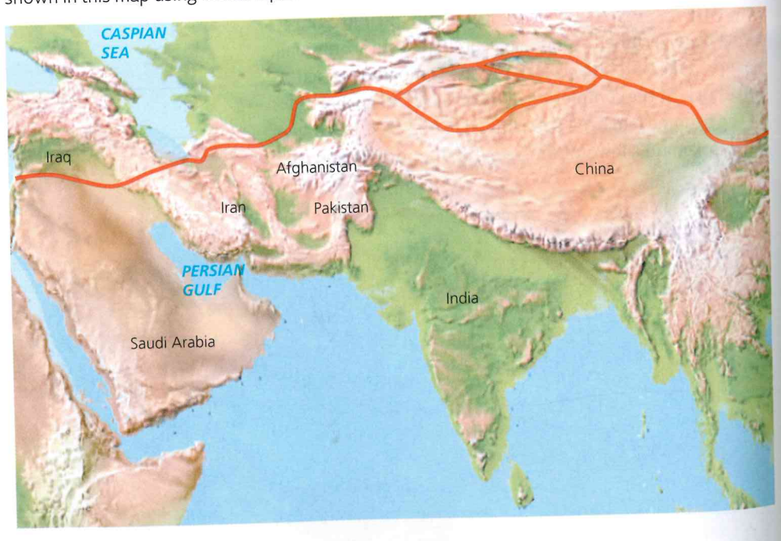
Scientists have even begun to gather sequence data from our extinct relatives. Neanderthals (*Homo neanderthalensis*) appeared at least 300,000 years ago in Europe and Asia and survived until about 30,000 years ago. Modern humans (*Homo sapiens*) first appeared in Africa around 200,000 years ago and spread into Europe and Asia around 50,000 years ago—meaning that modern humans and Neanderthals most likely comingled for a long time.

In 2013, scientists sequenced the entire genome of a 130,000-year-old female Neanderthal (*Homo neanderthalensis*). Using DNA extracted from a toe bone found in a Siberian cave, the resulting genome was nearly as complete as that from a modern human. Analysis of the Neanderthal genome revealed evidence of interbreeding with *Homo sapiens*. A 2014 study provided evidence that many present-day humans of European and Asian descent (but not African descent) carry Neanderthal-derived genes that influence the production of keratin, a protein that is a key structural component of hair, nails, and skin. Modern humans appear to have inherited the gene from Neanderthals around 70,000 years ago and then passed it on to their descendants. Such studies provide valuable insight into our own evolutionary tree.

Bioinformatics can also provide insights into our evolutionary relationships with nonhuman animals. In 2005, researchers completed the genome sequence for our closest living relative on the evolutionary tree of life, the chimpanzee (*Pan troglodytes*). Comparisons with human DNA revealed that we share 96% of our genome. Genomic scientists are currently finding and studying the important differences, shedding scientific light on the age-old question of what makes us human. **By comparing humans with related species both living (chimpanzees) and extinct (Neanderthals), researchers are shedding light on the recent evolutionary history of our own species.**

The potential benefits of knowing many human genomes are enormous. Thus far, more than 2,000 disease-associated genes have been identified. A recent example involved Behcet's disease, a painful and life-threatening illness that involves swelling of blood vessels throughout the body. Researchers have long known that this disease is found most commonly among people living along the ancient trade route in Asia called the Silk Road **(Figure 12.22)**. In 2013, researchers conducted a genome-wide search for genetic differences among Turkish people with and without the disease. They discovered four regions of the genome that are associated with the disease. The nearby genes are implicated in the immune system's ability to destroy invading microorganisms, to recognize infection sites, and to regulate autoimmune diseases. Interestingly, the function of the fourth gene has never been identified, but its close association with Behcet's disease may help researchers pinpoint its role. Next, in the Process of Science section, we'll examine how genomic analysis solved a medical mystery.

▼ Figure 12.22 The Silk Road. Behcet's disease is most commonly found along the Silk Road, part of which is shown in this map using modern place names.



Did Nic Have a Deadly Gene?

BACKGROUND

When infant Nic Volker stopped breast feeding, he became a medical mystery. Although previously healthy, he now cried in agony after meals and began to waste away. He developed ulcers between his intestines and skin and had to be fed through a nasal tube. After two years of mistaken diagnoses and 100 surgeries, everyone was desperate. Doctors hypothesized that Nic had a rare mutation that made his immune system attack his digestive system. But how could they prove that idea?

METHOD

A team of doctors at Medical College of Wisconsin sequenced Nic's entire protein-coding genome, a radical and expensive procedure in 2009. When doctors compared Nic's genome with other sequenced human genomes, they found over 16,000 mutations, the vast majority of which were not medically relevant. The challenge was to eliminate all mutations except for the one that was causing the symptoms. They developed new software to filter the mutations, eliminating ones that did not lead to malfunctioning proteins or were not involved in the digestion or immune systems. Then they reduced the list to rare mutations.

RESULTS

After eliminating thousands of mutations, the *XIAP* gene remained. This gene consists of about 500 nucleotides that code for the XIAP protein. The doctors found that Nic had a single base substitution, which produced a tyrosine amino acid where cysteine should be in the protein. In this case, as in many others, researchers gained insights

Applied Genomics

Solving the mystery of Nic Volker's rare mutation is a good example of applied genomics. In another medical application, a 2013 study used DNA sequencing to prove that cancerous skin cells that had spread to the brain had done so after fusing with red blood cells provided by a bone marrow donor. These results provided researchers with new insight into how cancer spreads throughout the body. Sequence data also provided strong evidence that a Florida dentist transmitted HIV to several patients and that a single natural strain of West Nile virus can infect both birds and people.

Applied genomics can also be used to investigate criminal cases. In 2001, a 63-year-old Florida man died from inhalation anthrax, a disease caused by breathing

of human DNA is of this type. Some of this noncoding DNA is made up of gene control sequences such as promoters, enhancers, and microRNAs (see Chapter 11). Other noncoding regions include introns and repetitive DNA (some of which is used in DNA profiling). Some noncoding DNA is important to our health, with certain regions known to carry disease-causing mutations. But the function of most noncoding DNA remains unknown.

ABOUT NUMBER AS A D ONLY ANY AS NT.

The human genome sequenced by government-funded scientists was actually a reference genome compiled from a group of individuals. As of today, the complete genomes of many individuals have been completed. Whereas sequencing the first human genome took 13 years and cost $100 million, we are rapidly approaching the day when an individual's genome can be sequenced in a matter of hours for less than $1,000.

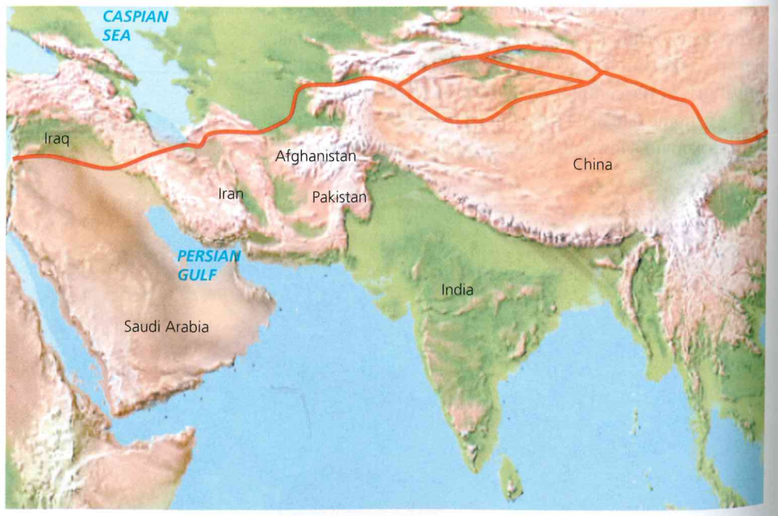
Scientists have even begun to gather sequence data from our extinct relatives. Neanderthals (*Homo neanderthalensis*) appeared at least 300,000 years ago in Europe and Asia and survived until about 30,000 years ago. Modern humans (*Homo sapiens*) first appeared in Africa around 200,000 years ago and spread into Europe and Asia around 50,000 years ago—meaning that modern humans and Neanderthals most likely comingled for a long time.

In 2013, scientists sequenced the entire genome of a 130,000-year-old female Neanderthal (*Homo neanderthalensis*). Using DNA extracted from a toe bone found in a Siberian cave, the resulting genome was nearly as complete as that from a modern human. Analysis of the Neanderthal genome revealed evidence of interbreeding with *Homo sapiens*. A 2014 study provided evidence that many present-day humans of European and Asian descent (but not African descent) carry Neanderthal-derived genes that influence the production of keratin, a protein that is a key structural component of hair, nails, and skin. Modern humans appear to have inherited the gene from Neanderthals around 70,000 years ago and then passed it on to their descendants. Such studies provide valuable insight into our own evolutionary tree.

Bioinformatics can also provide insights into our evolutionary relationships with nonhuman animals. In 2005, researchers completed the genome sequence for our closest living relative on the evolutionary tree of life, the chimpanzee (*Pan troglodytes*). Comparisons with human DNA revealed that we share 96% of our genome. Genomic scientists are currently finding and studying the important differences, shedding scientific light on the age-old question of what makes us human. **By comparing humans with related species both living (chimpanzees) and extinct (Neanderthals), researchers are shedding light on the recent evolutionary history of our own species.**

The potential benefits of knowing many human genomes are enormous. Thus far, more than 2,000 disease-associated genes have been identified. A recent example involved Behcet's disease, a painful and life-threatening illness that involves swelling of blood vessels throughout the body. Researchers have long known that this disease is found most commonly among people living along the ancient trade route in Asia called the Silk Road **(Figure 12.22)**. In 2013, researchers conducted a genome-wide search for genetic differences among Turkish people with and without the disease. They discovered four regions of the genome that are associated with the disease. The nearby genes are implicated in the immune system's ability to destroy invading microorganisms, to recognize infection sites, and to regulate autoimmune diseases. Interestingly, the function of the fourth gene has never been identified, but its close association with Behcet's disease may help researchers pinpoint its role. Next, in the Process of Science section, we'll examine how genomic analysis solved a medical mystery.

▼ Figure 12.22 The Silk Road. Behcet's disease is most commonly found along the Silk Road, part of which is shown in this map using modern place names.



THE PROCESS OF SCIENCE | DNA Profiling

BIOINFORMATICS

Did Nic Have a Deadly Gene?

BACKGROUND

When infant Nic Volker stopped breast feeding, he became a medical mystery. Although previously healthy, he now cried in agony after meals and began to waste away. He developed ulcers between his intestines and skin and had to be fed through a nasal tube. After two years of mistaken diagnoses and 100 surgeries, everyone was desperate. Doctors hypothesized that Nic had a rare mutation that made his immune system attack his digestive system. But how could they prove that idea?

METHOD

A team of doctors at Medical College of Wisconsin sequenced Nic's entire protein-coding genome, a radical and expensive procedure in 2009. When doctors compared Nic's genome with other sequenced human genomes, they found over 16,000 mutations, the vast majority of which were not medically relevant. The challenge was to eliminate all mutations except for the one that was causing the symptoms. They developed new software to filter the mutations, eliminating ones that did not lead to malfunctioning proteins or were not involved in the digestion or immune systems. Then they reduced the list to rare mutations.

RESULTS

After eliminating thousands of mutations, the *XIAP* gene remained. This gene consists of about 500 nucleotides that code for the XIAP protein. The doctors found that Nic had a single base substitution, which produced a tyrosine amino acid where cysteine should be in the protein. In this case, as in many others, researchers gained insights through comparison with other animals. Data from several species confirmed that cysteine is always at that key location on the XIAP protein **(Figure 12.23)**. The diagnosis allowed the team to save Nic's life by replacing his bone marrow with donor cells that contained functional *XIAP* genes. Today, similar procedures are using personal genome sequencing to save the lives of many children with rare mutations.

Nic Volker with a DNA-sequencing machine

▼ Figure 12.23 The sequence of amino acids in one region of the XIAP protein. This table shows the sequence of amino acids (using one-letter abbreviations) in the XIAP protein in patient Nic and other organisms. At a key location (highlighted on the top line), Nic had a base substitution that resulted in tyrosine (symbol: Y) instead of cysteine.

	Amino acids (one-letter abbreviations)																	
Nic	G	D	Q	V	Q	C	F	C	**Y**	G	G	K	L	K	N	W	E
Healthy human	G	D	Q	V	Q	C	F	C	C	G	G	K	L	K	N	W	E
Chicken	D	D	Q	V	Q	A	F	C	C	G	G	K	L	K	N	W	E
Zebra fish	D	D	N	V	Q	C	F	C	C	G	G	K	L	S	G	W	E
Frog	R	D	H	V	K	C	F	H	C	D	G	G	L	R	N	W	E
Housefly	L	D	H	V	K	C	V	W	C	N	G	V	I	A	K	W	E

Thinking Like a Scientist

Why would the scientists compare Nic's *XIAP* gene with genes in animals like flies and frogs?

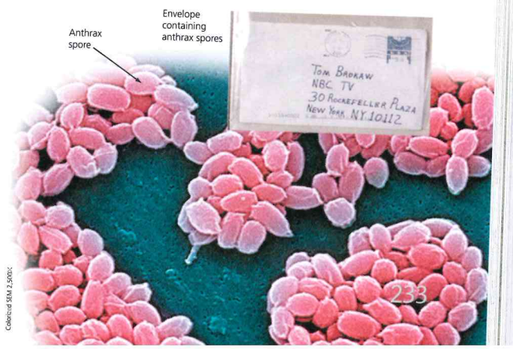
For the answer, see Appendix D.

Applied Genomics

Solving the mystery of Nic Volker's rare mutation is a good example of applied genomics. In another medical application, a 2013 study used DNA sequencing to prove that cancerous skin cells that had spread to the brain had done so after fusing with red blood cells provided by a bone marrow donor. These results provided researchers with new insight into how cancer spreads throughout the body. Sequence data also provided strong evidence that a Florida dentist transmitted HIV to several patients and that a single natural strain of West Nile virus can infect both birds and people.

Applied genomics can also be used to investigate criminal cases. In 2001, a 63-year-old Florida man died from inhalation anthrax, a disease caused by breathing spores of the bacterium *Bacillus anthracis*. Because he was the first victim of this disease in the United States since 1976 (and coming so soon after the 9/11 terrorist attacks the month before), his death was immediately considered suspicious. By the end of the year, four more people had died from inhaling anthrax. Law enforcement officials realized that someone was sending anthrax spores through the mail **(Figure 12.24)**. The United States was facing an unprecedented bio-terrorist attack.

▶ Figure 12.24 The 2001 anthrax attacks. In 2001, some envelopes containing anthrax spores caused five deaths.



ıe investigation that followed, one of the most help-
es turned out to be the anthrax spores themselves.
gators sequenced the genomes of the mailed anthrax
. They quickly established that all of the mailed
were genetically identical to a laboratory subtype
in a single flask at the U.S. Army Medical Research
ıte of Infectious Diseases in Fort Detrick, Maryland.
in part on this evidence, the FBI named an army
ch scientist as a suspect in the case. Although never
ed, the suspect committed suicide in 2008, and the
fficially remains unsolved.

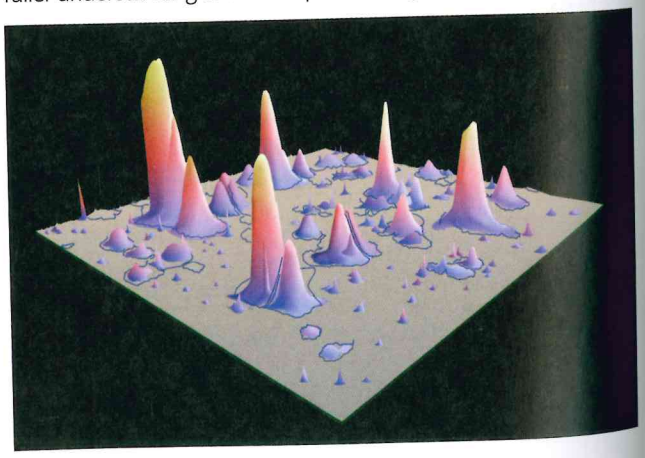
stems Biology

ɔmputational power provided by the tools of bioinfor-
s allows the study of whole sets of genes and their inter-
ıs, as well as the comparison of genomes from different
es. Genomics is a rich source of new insights into funda-
al questions about genome organization, regulation of
expression, embryonic development, and evolution.
chnological advances have also led to metagenomics,
udy of DNA from an environmental sample. When
ned directly from the environment, a sample will
contain genomes from many species. After the whole
ɔle is sequenced, computer software sorts out the partial
ences from different species and assembles them into
ıdividual specific genomes. So far, this approach has
applied to microbial communities found in environ-
ts as diverse as the Sargasso Sea and the human intes-
A 2012 study cataloged the astounding diversity of the
an "microbiome"—the many species of bacteria that
ist within and upon our bodies and that contribute to
survival. The ability to sequence the DNA of mixed popu-
ns eliminates the need to culture each species separately
ıe lab, making it more efficient to study microbial species.
'he successes in genomics have encouraged scientists
egin similar systematic studies of the full protein sets
genomes encode (proteomes), an approach called
teomics (Figure 12.25). The number of different pro-
s in humans far exceeds the number of different genes
ɔut 100,000 proteins versus about 21,000 genes). And
ause proteins, not genes, actually carry out the activi-
of the cell, scientists must study when and where pro-
ıs are produced and how they interact to understand
functioning of cells and organisms.
Genomics and proteomics enable biologists to approach
study of life from an increasingly global perspec-
e. Biologists are now compiling catalogs of genes and
ɔteins—that is, listings of all the "parts" that contribute
:he operation of cells, tissues, and organisms. As such
alogs become complete, researchers are shifting their
ention from the individual parts to how these parts
rk together in biological systems. This approach, called **systems biology**, aims to model the dynamic behavior of whole biological systems based on the study of the interactions among the system's parts. Because of the vast amounts of data generated in these types of studies, advances in computer technology and bioinformatics have been crucial in making systems biology possible.

Such analyses may have many practical applications. For example, proteins associated with specific diseases may be used to aid diagnosis (by developing tests that search for a particular combination of proteins) and treatment (by designing drugs that interact with the proteins involved). As high-throughput techniques become more rapid and less expensive, they are increasingly being applied to the problem of cancer. The Cancer Genome Atlas project is a simultaneous investigation by multiple research teams of a large group of interacting genes and gene products. This project aims to determine how changes in biological systems lead to cancer. A three-year pilot project set out to find all the common mutations in three types of cancer—lung, ovarian, and brain—by comparing gene sequences and patterns of gene expression in cancer cells with those in normal cells. The results confirmed the role of several genes suspected to be linked to cancer and identified a few previously unknown ones, suggesting possible new targets for therapies. The research approach proved so fruitful for these three types of cancer that the project has been extended to ten other types of cancer, chosen because they are common and often lethal in humans.

Systems biology is a very efficient way to study emergent properties, novel properties that arise at each successive level of biological complexity as a result of the arrangement of building blocks at the underlying level. **The more we can learn about the arrangement and interactions of the components of genetic systems, the deeper will be our understanding of whole organisms.** ✓

▼ **Figure 12.25 Proteomics.** Each peak on this three-dimensional graph represents one protein separated by gel electrophoresis. The height of the peak correlates with the amount of that protein. By identifying every protein in a sample, researchers can gain a fuller understanding of the complete biological system.

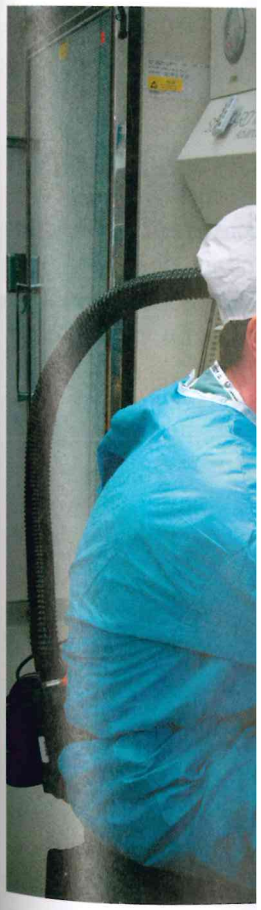


Safety a

As soon as scientists re
ogy, they began to wor
concerns focused on th
new disease-causing o
instance, if cancer-cau
infectious bacteria or v
scientists developed a
formal government reg
some other nations.

One safety measure
procedures to protect r
neered microorganism
from accidentally leavi
In addition, strains of r
recombinant DNA expe
ensure that they canno
a further precaution, c
ments have been banne

▼ **Figure 12.26 Maximu**
at the Centers for Diseas
he studies the virus that
killed over 50 million peo



In the investigation that followed, one of the most helpful clues turned out to be the anthrax spores themselves. Investigators sequenced the genomes of the mailed anthrax spores. They quickly established that all of the mailed spores were genetically identical to a laboratory subtype stored in a single flask at the U.S. Army Medical Research Institute of Infectious Diseases in Fort Detrick, Maryland. Based in part on this evidence, the FBI named an army research scientist as a suspect in the case. Although never charged, the suspect committed suicide in 2008, and the case officially remains unsolved.

Systems Biology

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

Technological advances have also led to metagenomics, the study of DNA from an environmental sample. When obtained directly from the environment, a sample will likely contain genomes from many species. After the whole sample is sequenced, computer software sorts out the partial sequences from different species and assembles them into the individual specific genomes. So far, this approach has been applied to microbial communities found in environments as diverse as the Sargasso Sea and the human intestine. A 2012 study cataloged the astounding diversity of the human "microbiome"—the many species of bacteria that coexist within and upon our bodies and that contribute to our survival. The ability to sequence the DNA of mixed populations eliminates the need to culture each species separately in the lab, making it more efficient to study microbial species.

The successes in genomics have encouraged scientists to begin similar systematic studies of the full protein sets that genomes encode (proteomes), an approach called **proteomics** (Figure 12.25). The number of different proteins in humans far exceeds the number of different genes (about 100,000 proteins versus about 21,000 genes). And because proteins, not genes, actually carry out the activities of the cell, scientists must study when and where proteins are produced and how they interact to understand the functioning of cells and organisms.

Genomics and proteomics enable biologists to approach the study of life from an increasingly global perspective. Biologists are now compiling catalogs of genes and proteins—that is, listings of all the "parts" that contribute to the operation of cells, tissues, and organisms. As such catalogs become complete, researchers are shifting their attention from the individual parts to how these parts work together in biological systems. This approach, called **systems biology**, aims to model the dynamic behavior of whole biological systems based on the study of the interactions among the system's parts. Because of the vast amounts of data generated in these types of studies, advances in computer technology and bioinformatics have been crucial in making systems biology possible.

Such analyses may have many practical applications. For example, proteins associated with specific diseases may be used to aid diagnosis (by developing tests that search for a particular combination of proteins) and treatment (by designing drugs that interact with the proteins involved). As high-throughput techniques become more rapid and less expensive, they are increasingly being applied to the problem of cancer. The Cancer Genome Atlas project is a simultaneous investigation by multiple research teams of a large group of interacting genes and gene products. This project aims to determine how changes in biological systems lead to cancer. A three-year pilot project set out to find all the common mutations in three types of cancer—lung, ovarian, and brain—by comparing gene sequences and patterns of gene expression in cancer cells with those in normal cells. The results confirmed the role of several genes suspected to be linked to cancer and identified a few previously unknown ones, suggesting possible new targets for therapies. The research approach proved so fruitful for these three types of cancer that the project has been extended to ten other types of cancer, chosen because they are common and often lethal in humans.

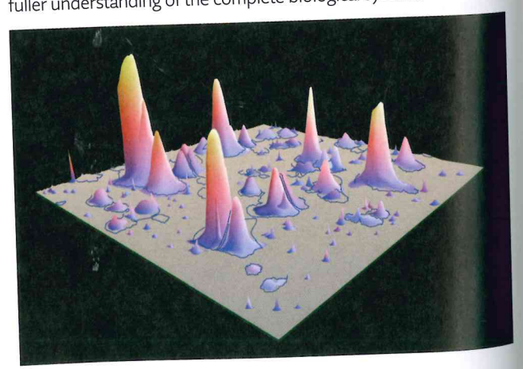
Systems biology is a very efficient way to study emergent properties, novel properties that arise at each successive level of biological complexity as a result of the arrangement of building blocks at the underlying level. **The more we can learn about the arrangement and interactions of the components of genetic systems, the deeper will be our understanding of whole organisms.** ✓

✓ CHECKPOINT

What is the difference between genomics and proteomics?

■ Answer: Genomics concerns the complete set of an organism's genes, whereas proteomics concerns the complete set of an organism's proteins.

▼ **Figure 12.25 Proteomics.** Each peak on this three-dimensional graph represents one protein separated by gel electrophoresis. The height of the peak correlates with the amount of that protein. By identifying every protein in a sample, researchers can gain a fuller understanding of the complete biological system.



Safety and Ethical Issues

As soon as scientists realized the power of DNA technology, they began to worry about potential dangers. Early concerns focused on the possibility of creating hazardous new disease-causing organisms. What might happen, for instance, if cancer-causing genes were transferred into infectious bacteria or viruses? To address such concerns, scientists developed a set of guidelines that have become formal government regulations in the United States and some other nations.

One safety measure in place is a set of strict laboratory procedures to protect researchers from infection by engineered microorganisms and to prevent microorganisms from accidentally leaving the laboratory (Figure 12.26). In addition, strains of microorganisms to be used in recombinant DNA experiments are genetically crippled to ensure that they cannot survive outside the laboratory. As a further precaution, certain obviously dangerous experiments have been banned.

▼ **Figure 12.26 Maximum-security laboratory.** A scientist at the Centers for Disease Control wears a biohazard suit as he studies the virus that caused the 1918 flu pandemic, which killed over 50 million people worldwide.

The C
Genet

Today, most
centers on g
account for
crops in the
together, th
the world's
crops is ofte
improve yie
for GM food
increasingly
must be gro

Controve
significant
European U
the market
the import
other natio
ceeded relat
foods is now

Advocate
crops carryi
environmer
example, in
cause allerg

▼ Figure 12.

In the investigation that followed, one of the most helpful clues turned out to be the anthrax spores themselves. Investigators sequenced the genomes of the mailed anthrax spores. They quickly established that all of the mailed spores were genetically identical to a laboratory subtype stored in a single flask at the U.S. Army Medical Research Institute of Infectious Diseases in Fort Detrick, Maryland. Based in part on this evidence, the FBI named an army research scientist as a suspect in the case. Although never charged, the suspect committed suicide in 2008, and the case officially remains unsolved.

Systems Biology

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

Technological advances have also led to metagenomics, the study of DNA from an environmental sample. When obtained directly from the environment, a sample will likely contain genomes from many species. After the whole sample is sequenced, computer software sorts out the partial sequences from different species and assembles them into the individual specific genomes. So far, this approach has been applied to microbial communities found in environments as diverse as the Sargasso Sea and the human intestine. A 2012 study cataloged the astounding diversity of the human "microbiome"—the many species of bacteria that coexist within and upon our bodies and that contribute to our survival. The ability to sequence the DNA of mixed populations eliminates the need to culture each species separately in the lab, making it more efficient to study microbial species.

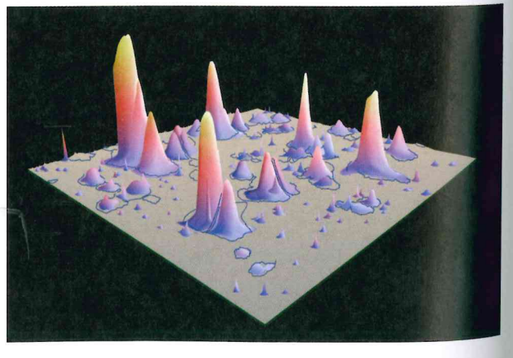
The successes in genomics have encouraged scientists to begin similar systematic studies of the full protein sets that genomes encode (proteomes), an approach called **proteomics** (Figure 12.25). The number of different proteins in humans far exceeds the number of different genes (about 100,000 proteins versus about 21,000 genes). And because proteins, not genes, actually carry out the activities of the cell, scientists must study when and where proteins are produced and how they interact to understand the functioning of cells and organisms.

Genomics and proteomics enable biologists to approach the study of life from an increasingly global perspective. Biologists are now compiling catalogs of genes and proteins—that is, listings of all the "parts" that contribute to the operation of cells, tissues, and organisms. As such catalogs become complete, researchers are shifting their attention from the individual parts to how these parts work together in biological systems. This approach, called **systems biology**, aims to model the dynamic behavior of whole biological systems based on the study of the interactions among the system's parts. Because of the vast amounts of data generated in these types of studies, advances in computer technology and bioinformatics have been crucial in making systems biology possible.

Such analyses may have many practical applications. For example, proteins associated with specific diseases may be used to aid diagnosis (by developing tests that search for a particular combination of proteins) and treatment (by designing drugs that interact with the proteins involved). As high-throughput techniques become more rapid and less expensive, they are increasingly being applied to the problem of cancer. The Cancer Genome Atlas project is a simultaneous investigation by multiple research teams of a large group of interacting genes and gene products. This project aims to determine how changes in biological systems lead to cancer. A three-year pilot project set out to find all the common mutations in three types of cancer—lung, ovarian, and brain—by comparing gene sequences and patterns of gene expression in cancer cells with those in normal cells. The results confirmed the role of several genes suspected to be linked to cancer and identified a few previously unknown ones, suggesting possible new targets for therapies. The research approach proved so fruitful for these three types of cancer that the project has been extended to ten other types of cancer, chosen because they are common and often lethal in humans.

Systems biology is a very efficient way to study emergent properties, novel properties that arise at each successive level of biological complexity as a result of the arrangement of building blocks at the underlying level. **The more we can learn about the arrangement and interactions of the components of genetic systems, the deeper will be our understanding of whole organisms.** ✓

▼ **Figure 12.25 Proteomics.** Each peak on this three-dimensional graph represents one protein separated by gel electrophoresis. The height of the peak correlates with the amount of that protein. By identifying every protein in a sample, researchers can gain a fuller understanding of the complete biological system.



Safety and Ethical Issues

As soon as scientists realized the power of DNA technology, they began to worry about potential dangers. Early concerns focused on the possibility of creating hazardous new disease-causing organisms. What might happen, for instance, if cancer-causing genes were transferred into infectious bacteria or viruses? To address such concerns, scientists developed a set of guidelines that have become formal government regulations in the United States and some other nations.

One safety measure in place is a set of strict laboratory procedures to protect researchers from infection by engineered microorganisms and to prevent microorganisms from accidentally leaving the laboratory (Figure 12.26). In addition, strains of microorganisms to be used in recombinant DNA experiments are genetically crippled to ensure that they cannot survive outside the laboratory. As a further precaution, certain obviously dangerous experiments have been banned.

▼ **Figure 12.26 Maximum-security laboratory.** A scientist at the Centers for Disease Control wears a biohazard suit as he studies the virus that caused the 1918 flu pandemic, which killed over 50 million people worldwide.

The Controversy over Genetically Modified Foods

Today, most public concern about possible hazards centers on genetically modified (GM) foods. GM strains account for a significant percentage of several staple crops in the United States, Argentina, and Brazil; together, these nations account for more than 80% of the world's supply of GM crops. The production of GM crops is often more profitable because the modifications improve yields and increase the size of harvest. Advocates for GM foods point out that transgenic crops will become increasingly necessary to combat starvation as more food must be grown on less space.

Controversy about the safety of these foods is a significant political issue (Figure 12.27). For example, the European Union has suspended the introduction into the market of new GM crops and considered banning the import of all GM foodstuffs. In the United States and other nations where the GM revolution initially proceeded relatively unnoticed, mandatory labeling of GM foods is now being debated.

Advocates of a cautious approach are concerned that crops carrying genes from other species might harm the environment or be hazardous to human health (by, for example, introducing new allergens, molecules that can cause allergic reactions, into foods). A major worry is that

▼ **Figure 12.27 Opposition to genetically modified organisms (GMOs).** Protesters in Oregon voice their displeasure about GMOs.

transgenic plants might pass their new genes to close relatives in nearby wild areas. We know that lawn and crop grasses, for example, commonly exchange genes with wild relatives through pollen transfer. If domestic plants carrying genes for resistance to herbicides, diseases, or insect pests pollinated wild plants, the offspring might become "superweeds" that would be very difficult to control. However, researchers may be able to prevent the escape of such plant genes in various ways—for example, by engineering plants so that they cannot breed. Concern has also been raised that the widespread use of GM seeds may reduce natural genetic diversity, leaving crops susceptible to catastrophic die-offs in the event of a sudden change to the environment or introduction of a new pest. Although the U.S. National Academy of Sciences released a study finding no scientific evidence that transgenic crops pose any special health or environmental risks, the authors of the study also recommended more stringent long-term monitoring to watch for unanticipated environmental impacts.

Negotiators from 130 nations (including the United States) agreed on a Biosafety Protocol that requires exporters to identify GM organisms present in bulk food shipments and allows importing nations to decide whether the shipments pose environmental or health risks. The United States declined to sign the agreement, but it went into effect anyway because the majority of nations were in favor of it. Since then, European nations have, on occasion, refused crops from the United States and other nations for fear that they contain GM crops, leading to trade disputes.

Governments and regulatory agencies throughout the world are grappling with how to facilitate the use of biotechnology in agriculture, industry, and medicine while ensuring that new products and procedures are safe. In the United States, all genetic engineering projects are evaluated for potential risks by a number of regulatory agencies, including the Food and Drug Administration, the Environmental Protection Agency, the National Institutes of Health, and the Department of Agriculture. ✅

✅ CHECKPOINT

What is the main concern about adding genes for herbicide resistance to crop plants?

■ Answer: the possibility that the genes could escape, through cross-pollination, to wild plants that are closely related to the crop species

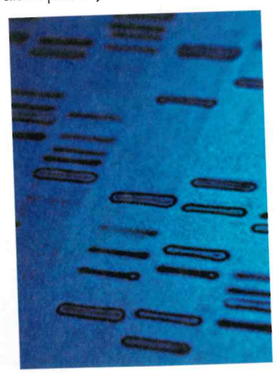
Ethical Questions Raised by Human DNA Technologies

Human DNA technologies raise legal and ethical questions—few of which have clear answers. Consider, for example, how the treatment of dwarfism with injections of human growth hormone (HGH) produced by genetically engineered cells might be extended beyond its current use. Should parents of short but hormonally normal children be able to seek HGH treatment to make their kids taller? If not, who decides which children are "tall enough" to be excluded from treatment? In addition to technical challenges, human gene therapy also provokes ethical questions. Some critics believe that tampering with human genes in any way is immoral or unethical. Other observers see no fundamental difference between the transplantation of genes into somatic cells and the transplantation of organs.

Genetic engineering of gametes (sperm or ova) and zygotes has been accomplished in lab animals. It has not been attempted in humans because such a procedure raises very difficult ethical questions. Should we try to eliminate genetic defects in our children and their descendants? Should we interfere with evolution in this way? From a long-term perspective, the elimination of unwanted versions of genes from the gene pool could backfire. **Genetic variety is a necessary ingredient for the adaptation of a species as environmental conditions change with time.** Genes that are damaging under some conditions may be advantageous under others (one example is the sickle-cell allele—see the Evolution Connection in Chapter 17). Are we willing to risk making genetic changes that could be detrimental to our species in the future?

Similarly, advances in genetic profiling raise privacy issues **(Figure 12.28)**. If we were to create a DNA profile of every person at birth, then theoretically we could match nearly every violent crime to a perpetrator because it is virtually impossible for someone to commit a violent crime without leaving behind DNA evidence. But are we, as a society, prepared to sacrifice our genetic privacy, even for such worthwhile goals? In 2014, the U.S. Supreme Court, by a 5–4 vote, upheld the practice of collecting DNA samples from suspects at the time of their arrest (before they had been convicted). Ruling that obtaining DNA is "like fingerprinting and photographing, a legitimate police booking procedure that is reasonable under the Fourth Amendment," the Supreme Court

▼ **Figure 12.28 Genetic information and privacy.** Collecting genomic data raises privacy concerns that affect everyone.



decision will likely usher in an era of ex
DNA profiling in many aspects of police

As more information becomes availal
personal genetic makeup, some people c
greater access to this information is alw
example, mail-in kits **(Figure 12.29)** have
that can tell healthy people their relativ
ing various diseases (such as Parkinson'
later in life. Some argue that such inforr
lies to prepare. Others worry that the te
fears without offering any real benefit b
diseases, such as Parkinson's, are not cu
able or treatable. Other tests, however, s
cancer risk, may help a person make cha
prevent disease. How can we identify tr

There is also a danger that informatic
associated genes could be abused. One is
sibility of discrimination and stigmatiza
the U.S. Congress passed the Genetic Inf
crimination Act of 2008. Title I of the ac
ance companies from requesting or requ
information during an application for h
Title II provides similar protections in e

A much broader ethical question is h
feel about wielding one of nature's powe
of new organisms? Some might ask if we

EVOLUTION CONNECTIO

The Y Chromosome

The human Y chromosome passes esse
father to son. Therefore, by comparing
ers can learn about the ancestry of hun
profiling can thus provide data about r
evolution.

Geneticists have discovered that ab
currently living in central Asia have Y
striking genetic similarity. Further an
common genetic heritage to a single m
1,000 years ago. In combination with h
the data led to the speculation that the
Genghis Khan **(Figure 12.30)** may have
for the spread of the chromosome to n
men living today. A similar study of Ir
gested that nearly 10% of them were d
Niall of the Nine Hostages, a warlord v
the 1400s. Another study of Y DNA se
the claim by the Lemba people of soutl
they are descended from ancient Jews
Sequences of Y DNA distinctive of the

transgenic plants might pass their new genes to close relatives in nearby wild areas. We know that lawn and crop grasses, for example, commonly exchange genes with wild relatives through pollen transfer. If domestic plants carrying genes for resistance to herbicides, diseases, or insect pests pollinated wild plants, the offspring might become "superweeds" that would be very difficult to control. However, researchers may be able to prevent the escape of such plant genes in various ways—for example, by engineering plants so that they cannot breed. Concern has also been raised that the widespread use of GM seeds may reduce natural genetic diversity, leaving crops susceptible to catastrophic die-offs in the event of a sudden change to the environment or introduction of a new pest. Although the U.S. National Academy of Sciences released a study finding no scientific evidence that transgenic crops pose any special health or environmental risks, the authors of the study also recommended more stringent long-term monitoring to watch for unanticipated environmental impacts.

Negotiators from 130 nations (including the United States) agreed on a Biosafety Protocol that requires exporters to identify GM organisms present in bulk food shipments and allows importing nations to decide whether the shipments pose environmental or health risks. The United States declined to sign the agreement, but it went into effect anyway because the majority of nations were in favor of it. Since then, European nations have, on occasion, refused crops from the United States and other nations for fear that they contain GM crops, leading to trade disputes.

Governments and regulatory agencies throughout the world are grappling with how to facilitate the use of biotechnology in agriculture, industry, and medicine while ensuring that new products and procedures are safe. In the United States, all genetic engineering projects are evaluated for potential risks by a number of regulatory agencies, including the Food and Drug Administration, the Environmental Protection Agency, the National Institutes of Health, and the Department of Agriculture. ✅

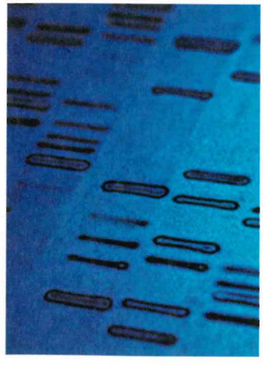
Ethical Questions Raised by Human DNA Technologies

Human DNA technologies raise legal and ethical questions—few of which have clear answers. Consider, for example, how the treatment of dwarfism with injections of human growth hormone (HGH) produced by genetically engineered cells might be extended beyond its current use. Should parents of short but hormonally normal children be able to seek HGH treatment to make their kids taller? If not, who decides which children are "tall enough" to be excluded from treatment? In addition to technical challenges, human gene therapy also provokes ethical questions. Some critics believe that tampering with human genes in any way is immoral or unethical. Other observers see no fundamental difference between the transplantation of genes into somatic cells and the transplantation of organs.

Genetic engineering of gametes (sperm or ova) and zygotes has been accomplished in lab animals. It has not been attempted in humans because such a procedure raises very difficult ethical questions. Should we try to eliminate genetic defects in our children and their descendants? Should we interfere with evolution in this way? From a long-term perspective, the elimination of unwanted versions of genes from the gene pool could backfire. **Genetic variety is a necessary ingredient for the adaptation of a species as environmental conditions change with time.** Genes that are damaging under some conditions may be advantageous under others (one example is the sickle-cell allele—see the Evolution Connection in Chapter 17). Are we willing to risk making genetic changes that could be detrimental to our species in the future?

Similarly, advances in genetic profiling raise privacy issues **(Figure 12.28)**. If we were to create a DNA profile of every person at birth, then theoretically we could match nearly every violent crime to a perpetrator because it is virtually impossible for someone to commit a violent crime without leaving behind DNA evidence. But are we, as a society, prepared to sacrifice our genetic privacy, even for such worthwhile goals? In 2014, the U.S. Supreme Court, by a 5–4 vote, upheld the practice of collecting DNA samples from suspects at the time of their arrest (before they had been convicted). Ruling that obtaining DNA is "like fingerprinting and photographing, a legitimate police booking procedure that is reasonable under the Fourth Amendment," the Supreme Court decision will likely usher in an era of expanded use of DNA profiling in many aspects of police work.

▼ **Figure 12.28 Genetic information and privacy.** Collecting genomic data raises privacy concerns that affect everyone.



As more information becomes available about our personal genetic makeup, some people question whether greater access to this information is always beneficial. For example, mail-in kits **(Figure 12.29)** have become available that can tell healthy people their relative risk of developing various diseases (such as Parkinson's and Crohn's) later in life. Some argue that such information helps families to prepare. Others worry that the tests prey on our fears without offering any real benefit because certain diseases, such as Parkinson's, are not currently preventable or treatable. Other tests, however, such as for breast cancer risk, may help a person make changes that can prevent disease. How can we identify truly useful tests?

There is also a danger that information about disease-associated genes could be abused. One issue is the possibility of discrimination and stigmatization. In response, the U.S. Congress passed the Genetic Information Nondiscrimination Act of 2008. Title I of the act prohibits insurance companies from requesting or requiring genetic information during an application for health insurance. Title II provides similar protections in employment.

A much broader ethical question is how do we really feel about wielding one of nature's powers—the evolution of new organisms? Some might ask if we have any right to alter an organism's genes—or to create new organisms. DNA technologies raise many complex issues that have no easy answers. It is up to you, as a participating citizen, to make informed choices. ✅

▼ **Figure 12.29 Personalized genetic testing.** This kit can be used to send saliva for genetic analysis. The results can indicate a person's risk of developing certain diseases.

✅ CHECKPOINT

Why does genetically modifying a human gamete raise different ethical questions than genetically modifying a human somatic (body) cell?

■ *Answer: A genetically modified somatic cell will affect only the patient. Modifying a gamete will affect an unborn individual as well as all of his or her descendants.*

EVOLUTION CONNECTION DNA Profiling

The Y Chromosome as a Window on History

The human Y chromosome passes essentially intact from father to son. Therefore, by comparing Y DNA, researchers can learn about the ancestry of human males. DNA profiling can thus provide data about recent human evolution.

Geneticists have discovered that about 8% of males currently living in central Asia have Y chromosomes of striking genetic similarity. Further analysis traced their common genetic heritage to a single man living about 1,000 years ago. In combination with historical records, the data led to the speculation that the Mongolian ruler Genghis Khan **(Figure 12.30)** may have been responsible for the spread of the chromosome to nearly 16 million men living today. A similar study of Irish men suggested that nearly 10% of them were descendants of Niall of the Nine Hostages, a warlord who lived during the 1400s. Another study of Y DNA seemed to confirm the claim by the Lemba people of southern Africa that they are descended from ancient Jews **(Figure 12.31)**. Sequences of Y DNA distinctive of the Jewish priestly caste called Kohanim are found at high frequencies among the Lemba.

Comparison of Y chromosome DNA profiles is part of a larger effort to learn more about the human genome. Other research efforts are extending genomic studies to many more species. These studies will advance our understanding of all aspects of biology, including health and ecology, as well as evolution. In fact, comparisons of the completed genome sequences of bacteria, archaea, and eukaryotes first supported the theory that these are the three fundamental domains of life—a topic we discuss further in the next unit, Evolution and Diversity.

▲ **Figure 12.30** Genghis Khan.

▶ **Figure 12.31 Lemba people of southern Africa.**

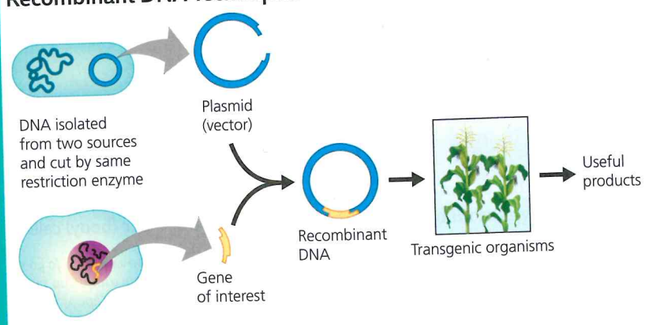
Chapter Review

SUMMARY OF KEY CONCEPTS

Genetic Engineering

DNA technology, the manipulation of genetic material, is a relatively new branch of biotechnology, the use of organisms to make helpful products. DNA technology often involves the use of recombinant DNA, the combination of nucleotide sequences from two different sources.

Recombinant DNA Techniques



Gene Editing

The CRISPR-Cas9 system can be used to deactivate or edit genes within living cells.

Medical Applications

By transferring a human gene into a bacterium or other easy-to-grow cell, scientists can mass-produce valuable human proteins to be used as drugs or vaccines.

Genetically Modified Organisms in Agriculture

Recombinant DNA techniques have been used to create genetically modified organisms, those that carry artificially introduced genes. Nonhuman cells have been engineered to produce human proteins, genetically modified food crops, and transgenic farm animals. A transgenic organism is one that carries artificially introduced genes, typically from a different species.

Human Gene Therapy

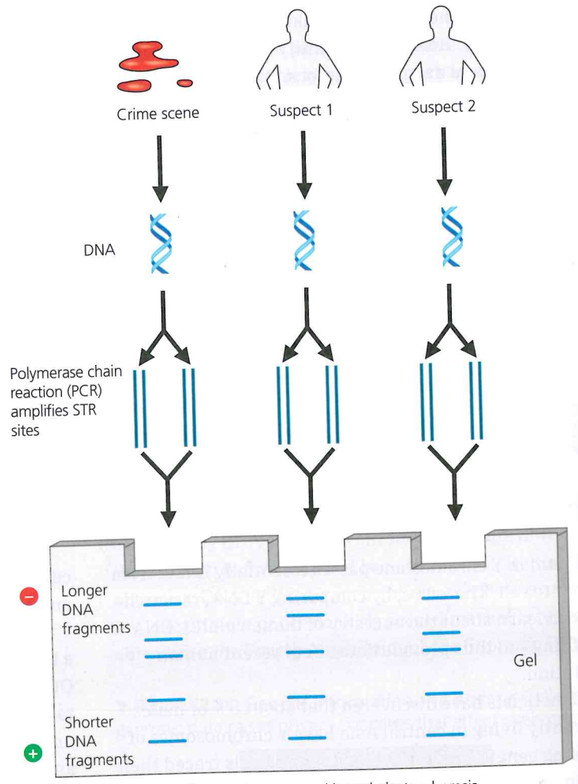
A virus can be modified to include a normal human gene. If this virus is injected into the bone marrow of a person suffering from a genetic disease, the normal human gene may be transcribed and translated, producing a normal human protein that may cure the genetic disease. This technique has been used in gene therapy trials involving a number of inherited diseases. There have been both successes and failures to date, and research continues.

DNA Profiling and Forensic Science

Forensics, the scientific analysis of legal evidence, has been revolutionized by DNA technology. DNA profiling is used to determine whether two DNA samples come from the same individual.

DNA Profiling Techniques

Short tandem repeat (STR) analysis compares DNA fragments using the polymerase chain reaction (PCR) and gel electrophoresis.



DNA fragments compared by gel electrophoresis
(Bands of shorter fragments move faster toward the positive pole.)

Investigating Murder, Paternity, and Ancient DNA

DNA profiling can be used to establish innocence or guilt of a criminal suspect, identify victims, determine paternity, and contribute to basic research.

Bioinformatics

DNA Sequencing

Automated machines can now sequence many thousands of DNA nucleotides per hour.

Genomics

Advances in DNA sequencing have ushered in the era of genomics, the study of complete genome sets.

Genome-Mapping Techniques

The whole-genome shotgun method involves sequencing DNA fragments from an entire genome and then assembling the sequences.

The Human Genome

The nucleotide sequence of the human genome is providing a wealth of useful data. The 24 different chromosomes of the human genome contain about 3 billion nucleotide pairs and 21,000 genes. The majority of the genome consists of noncoding DNA.

Applied Genomics

Comparing genomes can aid criminal investigations and basic biological research.

Systems Biology

Success in genomics has given rise to proteomics, the systematic study of the full set of proteins found in organisms. Genomics and proteomics both contribute to systems biology, the study of how many parts work together within complex biological systems.

Safety and Ethical Issues

The Controversy over Genetically Modified Foods

The debate about genetically modified crops centers on whether they might harm humans or damage the environment by transferring genes through cross-pollination with other species.

Ethical Questions Raised by Human DNA Technologies

As members of society we must become educated about DNA technologies so that we can intelligently address the ethical questions raised by their use.

Mastering Biology

For practice quizzes, BioFlix animations, MP3 tutorials, video tutors, and more study tools designed for this textbook, go to Mastering Biology™

SELF-QUIZ

1. Suppose you wish to create a recombinant DNA. Place the f have to perform them.
   a. Find the clone with the gen
   b. Insert the plasmids into bac into clones.
   c. Isolate the gene for lactase.
   d. Create recombinant plasmi for lactase.
2. A carrier that moves DNA fron called a _____.
3. In making recombinant DNA, enzyme that cuts DNA in a sta
4. A paleontologist has recovere 400-year-old preserved skin c compare DNA from the sampl most useful method for initial available for testing is _____
5. Why do DNA fragments conta to migrate to different locatio
6. What feature of a DNA fragme electrophoresis?
   a. the electrical charges of its
   b. its nucleotide sequence
   c. the hydrogen bonds betwe
   d. its double helix shape
7. After a gel electrophoresis pro gel shows
   a. the order of bases in a part
   b. the presence of various-siz
   c. the order of genes along pa
   d. the exact location of a spec
8. Name the steps of the whole-
9. Put the following steps of hum
   a. Virus is injected into patien
   b. Human gene is inserted int
   c. Normal human gene is isol
   d. Normal human gene is tran in the patient.

For answers to the Self Quiz, see Appena

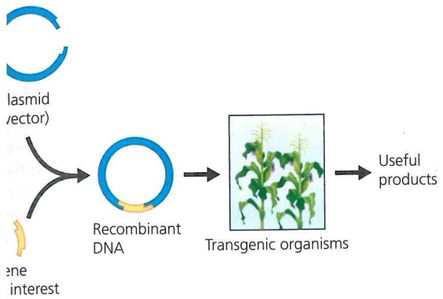
Review

KEY CONCEPTS

ring

nipulation of genetic material, is a relatively new
, the use of organisms to make helpful products.
nvolves the use of recombinant DNA, the combina-
nces from two different sources.

echniques



m can be used to deactivate or edit genes within

gene into a bacterium or other easy-to-grow cell,
duce valuable human proteins to be used as drugs or

l Organisms in Agriculture

niques have been used to create genetically modi-
at carry artificially introduced genes. Nonhuman
ered to produce human proteins, genetically modified
nic farm animals. A transgenic organism is one that
uced genes, typically from a different species.

oy

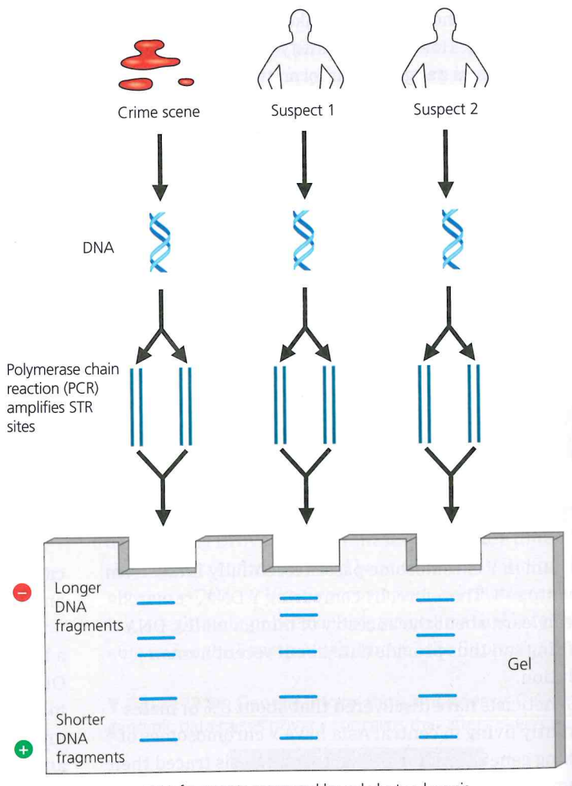
l to include a normal human gene. If this virus is
narrow of a person suffering from a genetic disease,
e may be transcribed and translated, producing a
that may cure the genetic disease. This technique
therapy trials involving a number of inherited dis-
both successes and failures to date, and research

DNA Profiling and Forensic Science

Forensics, the scientific analysis of legal evidence, has been revolutionized by DNA technology. DNA profiling is used to determine whether two DNA samples come from the same individual.

DNA Profiling Techniques

Short tandem repeat (STR) analysis compares DNA fragments using the polymerase chain reaction (PCR) and gel electrophoresis.



DNA fragments compared by gel electrophoresis
(Bands of shorter fragments move faster toward the positive pole.)

Investigating Murder, Paternity, and Ancient DNA

DNA profiling can be used to establish innocence or guilt of a criminal suspect, identify victims, determine paternity, and contribute to basic research.

Bioinformatics

DNA Sequencing

Automated machines can now sequence many thousands of DNA nucleotides per hour.

Genomics

Advances in DNA sequencing have ushered in the era of genomics, the study of complete genome sets.

Genome-Mapping Techniques

The whole-genome shotgun method involves sequencing DNA fragments from an entire genome and then assembling the sequences.

The Human Genome

The nucleotide sequence of the human genome is providing a wealth of useful data. The 24 different chromosomes of the human genome contain about 3 billion nucleotide pairs and 21,000 genes. The majority of the genome consists of noncoding DNA.

Applied Genomics

Comparing genomes can aid criminal investigations and basic biological research.

Systems Biology

Success in genomics has given rise to proteomics, the systematic study of the full set of proteins found in organisms. Genomics and proteomics both contribute to systems biology, the study of how many parts work together within complex biological systems.

Safety and Ethical Issues

The Controversy over Genetically Modified Foods

The debate about genetically modified crops centers on whether they might harm humans or damage the environment by transferring genes through cross-pollination with other species.

Ethical Questions Raised by Human DNA Technologies

As members of society we must become educated about DNA technologies so that we can intelligently address the ethical questions raised by their use.

Mastering Biology

For practice quizzes, BioFlix animations, MP3 tutorials, video tutors, and more study tools designed for this textbook, go to Mastering Biology™

SELF-QUIZ

1. Suppose you wish to create a large batch of the protein lactase using recombinant DNA. Place the following steps in the order you would have to perform them.
   a. Find the clone with the gene for lactase.
   b. Insert the plasmids into bacteria and grow the bacteria into clones.
   c. Isolate the gene for lactase.
   d. Create recombinant plasmids, including one that carries the gene for lactase.
2. A carrier that moves DNA from one cell to another, such as a plasmid, is called a _____.
3. In making recombinant DNA, what is the benefit of using a restriction enzyme that cuts DNA in a staggered fashion?
4. A paleontologist has recovered a bit of organic material from the 400-year-old preserved skin of an extinct dodo. She would like to compare DNA from the sample with DNA from living birds. The most useful method for initially increasing the amount of dodo DNA available for testing is _____.
5. Why do DNA fragments containing STR sites from different people tend to migrate to different locations during gel electrophoresis?
6. What feature of a DNA fragment causes it to move through a gel during electrophoresis?
   a. the electrical charges of its phosphate groups
   b. its nucleotide sequence
   c. the hydrogen bonds between its base pairs
   d. its double helix shape
7. After a gel electrophoresis procedure is run, the pattern of bars in the gel shows
   a. the order of bases in a particular gene.
   b. the presence of various-sized fragments of DNA.
   c. the order of genes along particular chromosomes.
   d. the exact location of a specific gene in the genome.
8. Name the steps of the whole-genome shotgun method.
9. Put the following steps of human gene therapy in the correct order.
   a. Virus is injected into patient.
   b. Human gene is inserted into a virus.
   c. Normal human gene is isolated and cloned.
   d. Normal human gene is transcribed and translated in the patient.

For answers to the Self Quiz, see Appendix D.