15

DNA Technology and Genomics

Keith V. Wood/Visuals Unlimited

Genetically engineered tobacco plant. This plant glows as it expresses the luciferase gene from a firefly. Luciferase is an enzyme that catalyzes a reaction that produces a flash of light. This classic experiment indicated that animal genes can be expressed in plants.

KEY CONCEPTS

Recombinant DNA techniques allow scientists to clone many copies of specific genes and gene products.

Biologists study DNA using gel electrophoresis, DNA blotting, automated sequencing, and other methods.

Genomics is an emerging field that comprises the structure, function, and evolution of genomes.

DNA technology and genomics have wide applications, from medical to forensic to agricultural.

Beginning in the mid-1970s, the development of new ways to study DNA led to radically new research approaches that have had a major impact in areas from cell biology and evolution to ethical and societal issues. This chapter begins with a consideration of **recombinant DNA technology,** in which researchers splice together DNA from different organisms in the laboratory. One goal of this technology is to enable scientists to obtain many copies of a specific DNA segment for the purpose of studying it. Because new methods to analyze DNA are continually emerging, we do not attempt to explore them all here. Instead, we discuss some of the major approaches that have provided a foundation for the technology.

We then consider how studies of DNA sequences have helped scientists understand the organization of genes and the relationship between genes and their products. In fact, most of our current knowledge of the structure and control of eukaryotic genes, and of the roles of genes in development, comes from applying these methods.

This chapter also explores some of the practical applications of DNA technologies. One of the rapidly advancing areas of study is **genetic engineering**—modifying the DNA of an organism to produce new genes with new traits. Genetic engineering can take many forms, ranging from basic research (see photograph) to the production of strains of bacteria that manufacture useful protein products, to the development of plants and animals that express foreign genes useful in agriculture.

The development of DNA cloning, genetic engineering, and related techniques has transformed people's view of **biotechnology,** the use of organisms to develop useful products. A traditional example of biotechnology is the use of yeast to make alcoholic beverages. Today, biotechnology includes numerous applications in such diverse areas as medicine and the pharmaceutical industry, foods and agriculture, and forensic science. ■

DNA CLONING

Learning Objectives

1 Explain how a typical restriction enzyme cuts DNA molecules, and give examples of the ways in which these enzymes are used in recombinant DNA technology.
2 Distinguish among a genomic DNA library, a chromosome library, and a complementary DNA (cDNA) library; explain why one would clone the same eukaryotic gene from both a genomic DNA library and a cDNA library.
3 Describe the purpose of a genetic probe.
4 Describe how the polymerase chain reaction amplifies DNA in vitro.

Recombinant DNA technology was not developed quickly. It actually had its roots in the 1940s with genetic studies of bacteria and **bacteriophages** ("bacteria eaters"), the viruses that infect them (see Fig. 12-2). After decades of basic research and the accumulation of extensive knowledge, the technology became feasible and available to the many scientists who now use these methods.

In recombinant DNA technology, scientists use enzymes from bacteria, known as **restriction enzymes,** to cut DNA molecules only in specific places. Restriction enzymes enable researchers to cut DNA into manageable segments. Each fragment is then incorporated into a suitable **vector** molecule, a carrier capable of transporting the DNA fragment into a cell. Bacteriophages and DNA molecules called plasmids are two examples of vectors. Bacterial DNA is circular; a **plasmid** is a separate, much smaller, circular DNA molecule that may be present and replicate inside a bacterial cell, such as *Escherichia coli.* Researchers introduce plasmids into bacterial cells by a method called **transformation,** the uptake of foreign DNA by cells (see Chapter 12). For transformation to be efficient, the researcher alters the cells chemically or by *electroporation*—delivering an electric shock—to make the plasma membrane permeable to the plasmid DNA molecules. Once a plasmid enters a cell, it is replicated and distributed to daughter cells during cell division. When a *recombinant plasmid*—one that has foreign DNA spliced into it—replicates in this way, many identical copies of the foreign DNA are made; in other words, the foreign DNA is **cloned.**

Restriction enzymes are "molecular scissors"

Discovering restriction enzymes was a major breakthrough in developing recombinant DNA technology. Today, large numbers of different types of restriction enzymes, each with its own characteristics, are readily available to researchers. For example, a restriction enzyme known as *Hin*dIII recognizes and cuts a DNA molecule at the restriction site 5′—AAGCTT—3′. (A *restriction site* is a DNA sequence containing the cleavage site that is cut by a particular restriction enzyme.) The sequence 5′—GAATTC—3′ is cut by another restriction enzyme, known as *Eco*RI. The names of restriction enzymes are generally derived from the names of the bacteria from which they were originally isolated. Thus, *Hin*dIII and *Eco*RI are derived from *Hemophilus influenzae* and *E. coli,* respectively.

Why do bacteria produce such enzymes? During infection, a bacteriophage injects its DNA into a bacterial cell. The bacterium can defend itself if it has restriction enzymes that can attack the bacteriophage DNA. The cell protects its own DNA from breakdown by modifying it after replication. An enzyme adds a methyl group to one or more bases in each restriction site so that the restriction enzyme does not recognize and cut the bacterial DNA.

Restriction enzymes enable scientists to cut DNA from chromosomes into shorter fragments in a controlled way. Many of the restriction enzymes used for recombinant DNA studies cut **palindromic sequences,** which means the base sequence of one strand reads the same as its complement when both are read in the 5′ ⟶ 3′ direction. Thus, in our *Hin*dIII example, both strands read 5′—AAGCTT—3′, which as a double-stranded molecule is diagrammed as follows:

```
5′—AAGCTT—3′
3′—TTCGAA—5′
```

By cutting both strands of the DNA, but in a staggered fashion, these enzymes produce fragments with identical, complementary, single-stranded ends:

```
5′—A        AGCTT—3′
3′—TTCGA        A—5′
```

These ends are called *sticky ends* because they pair by hydrogen bonding with the complementary, single-stranded ends of other DNA molecules that have been cut with the same enzyme (▮ Fig. 15-1). Once the sticky ends of two molecules have been joined in this way, they are treated with **DNA ligase,** an enzyme that covalently links the two DNA fragments to form a stable recombinant DNA molecule.

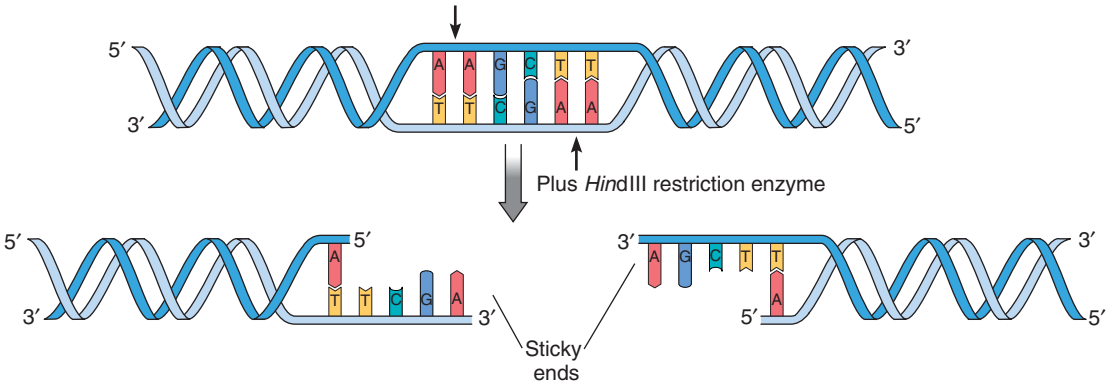

Figure 15-1 ***Animated*** Cutting DNA with a restriction enzyme

Many restriction enzymes, such as *Hin*dIII, cut DNA at sequences that are palindromic, producing complementary sticky ends. The small black arrows designate the enzyme's cleavage sites.

Recombinant DNA forms when DNA is spliced into a vector

In recombinant DNA technology, geneticists cut both foreign DNA and plasmid DNA with the same restriction enzyme. The two types of DNA are then mixed under conditions that facilitate hydrogen bonding between the complementary bases of the sticky ends, and the nicks in the resulting recombinant DNA are sealed by DNA ligase (Fig. 15-2).

The plasmids now used in recombinant DNA work have been extensively manipulated in the laboratory to include features helpful in isolating and analyzing cloned DNA (Fig. 15-3). Among these are an origin of replication (see Chapter 12), one or more restriction sites, and genes that let researchers select cells transformed by recombinant plasmids. These genes cause transformed cells to grow under specified conditions that do not allow untransformed cells to grow. In this way, the researchers use features that are also common in naturally occurring plasmids. Typically, plasmids do not contain genes essential to the bacterial cells under normal conditions. However, they often carry genes that are useful under specific environmental conditions, such as genes that confer resistance to particular antibiotics or that let the cells use a specific nutrient. For example, cells transformed with a plasmid that includes a gene for resistance to the antibiotic tetracycline can grow in a medium that contains tetracycline, whereas untransformed cells cannot.

A limiting property of any vector is the size of the DNA fragment it can effectively carry. The size of a DNA segment is often given in kilobases (kb), with 1 kb being equal to 1000 base pairs. Fragments that are smaller than 10 kb are usually inserted into plasmids for use in *E. coli.* However, larger fragments require the use of bacteriophage vectors, which can handle up to 23 kb of DNA. Other vectors are **cosmid cloning vectors,** which are combination vectors with features from both bacteriophages and plasmids, and **bacterial artificial chromosomes (BACs),** which accommodate much larger fragments of DNA. A BAC can include up to about 200 kb of extra DNA, a feature that made BACs especially useful in the *Human Genome Project,* which is discussed later in the chapter.

Recombinant DNA can also be introduced into cells of eukaryotic organisms. For example, geneticists use engineered viruses as vectors in mammal cells. These viruses are disabled so they do not kill the cells they infect. Instead, the viral DNA, as well as any foreign DNA they carry, becomes incorporated into the cell's chromosomes after infection. As discussed later, other methods do not require a biological vector.

DNA can be cloned inside cells

Because a single gene is only a small part of the total DNA in an organism, isolating the piece of DNA containing that particular gene is like finding a needle in a haystack: a powerful detector is needed. Today, many methods enable biologists to isolate a specific nucleotide sequence from an organism. We start by discussing methods in which DNA is cloned inside bacterial cells. We use the cloning of human DNA as an example, although the procedure is applicable for any organism.

A genomic DNA library contains fragments of all DNA in the genome

The total DNA in a cell is called its **genome.** For example, if DNA is extracted from human cells, we refer to it as human genomic DNA. A **genomic DNA library** is a collection of thousands of DNA fragments that represent all the DNA in the genome. Each fragment is inserted into a plasmid, which is usually incorporated into a bacterial cell. Thus, a human genomic DNA library is stored in a collection of recombinant bacteria, each with a different fragment of human DNA. Scientists use genomic DNA libraries to isolate and study specific genes.

Individual chromosomes can also be isolated to make a **chromosome library** containing all the DNA fragments in that

Key Point

Recombinant DNA technology involves a series of steps that make use of biochemistry, genetics, and molecular biology.

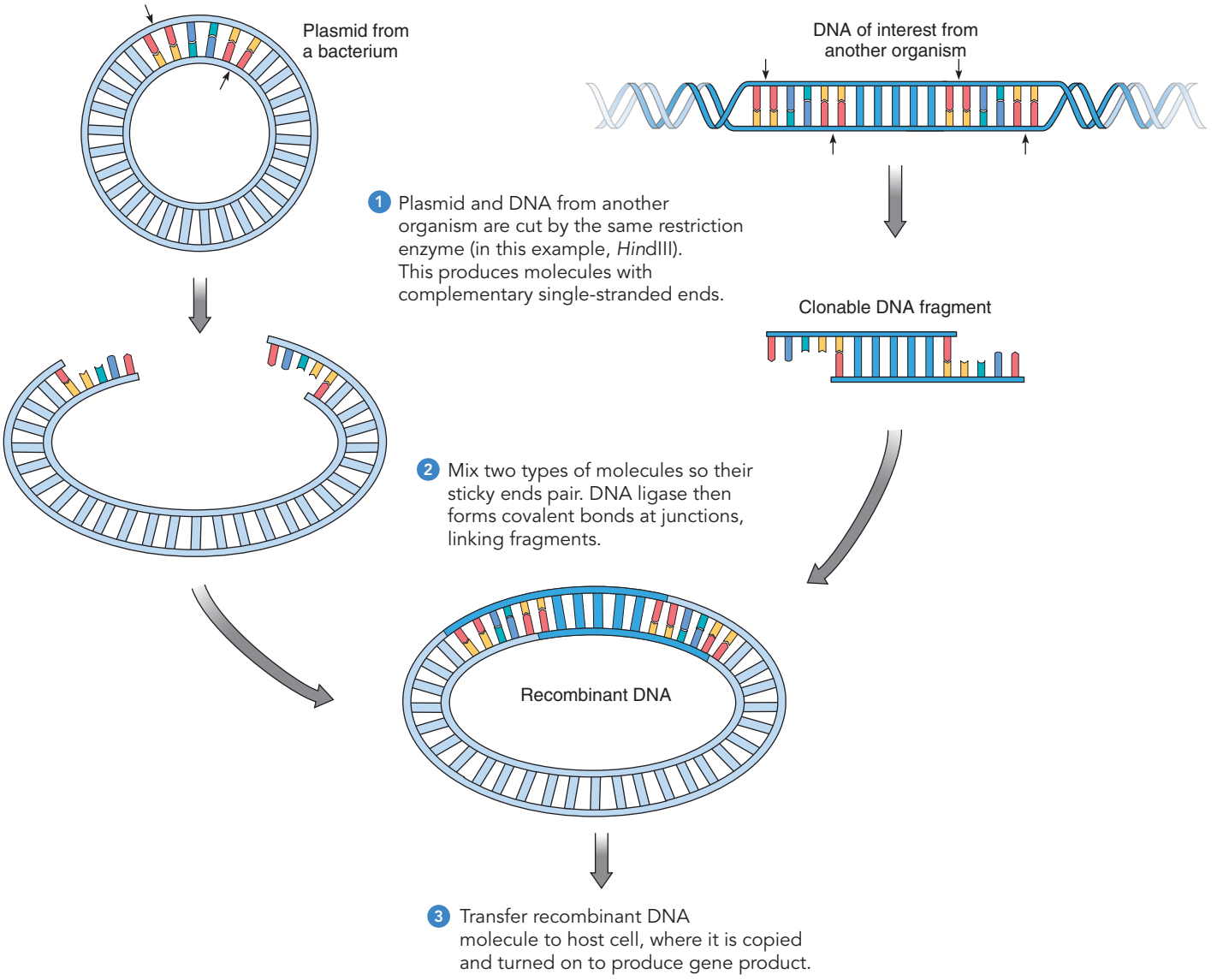

Figure 15-2 ***Animated*** Splicing foreign DNA into a vector

In this example, a bacterial plasmid is the vector.

specific chromosome. If a gene of interest is known to be associated with a particular chromosome, it is easier to isolate that gene from a chromosome library than from a genomic DNA library.

The basic cloning methods just described are used to make genomic DNA or chromosome libraries. First, the DNA is cut with a restriction enzyme, generating a population of DNA fragments (Fig. 15-4 1). These fragments vary in size and in the genetic information they carry, but they all have identical sticky ends. Geneticists treat plasmid DNA that will be used as a vector with the same restriction enzyme, which converts the circular plasmids into linear molecules with sticky ends complementary to those of the human DNA fragments. Recombinant plasmids are produced by mixing the two kinds of DNA (human and plasmid) under conditions that promote hydrogen bonding of complementary bases. Then, DNA ligase is used to covalently bond the paired ends of the plasmid and human DNA, forming recombinant DNA (Fig. 15-4 2). Unavoidably, nonrecombinant plasmids also form (*not shown in figure*), because some plasmids revert to their original circular shape without incorporating foreign DNA.

in a field of wheat. The transgenic weed was a poor competitor and had less effect on wheat production than its wild relatives in a control field. These results, although encouraging, must be interpreted with care. Scientists must evaluate each transgenic crop plant individually to see if there is gene flow to wild relatives, and if so, what effect might result.

Other concerns relate to plants engineered to produce pesticides, such as the *Bt* toxin. The future of the *Bt* toxin in transgenic crops is not secure, because low levels of the insecticide could potentially provide ideal conditions for selection for resistant individuals in the insect population. It appears certain insects may evolve genetic resistance to the *Bt* toxin in transgenic plants in the same way they evolve genetic resistance to chemical insecticides.

Another concern is that nonpest species could be harmed. For example, people paid a great deal of attention to the finding that monarch butterfly larvae raised in the laboratory are harmed if they are fed pollen from *Bt* corn plants. Although more recent studies suggest that monarch larvae living in a natural environment do not consume enough pollen to cause damage, such concerns persist and will have to be addressed individually.

Environmental concerns about transgenic animals also exist. Several countries are in the process of developing fast-growing transgenic fish, usually by inserting a gene that codes for a growth hormone. Transgenic Atlantic salmon, for example, grow up to 6 times as fast as nontransgenic salmon grown for human consumption. The transgenic fish do not grow larger than other fish, just faster. The benefits of such genetically enhanced fish include reduced pressure on wild fisheries and less pollution from fish farms. However, if the transgenic fish escaped from the fish farm, what effect would they have on wild relatives? To address this concern, the researchers are developing only nonreproducing female transgenic salmon.

To summarize, DNA technology in agriculture offers many potential benefits, including higher yields by providing disease resistance, more nutritious foods, and the reduced use of chemical pesticides. However, like other kinds of technology, genetic engineering poses some risks, such as the risk that genetically modified plants and animals could pass their foreign genes to wild relatives, causing unknown environmental problems. The science of **risk assessment,** which uses statistical methods to quantify risks so they can be compared and contrasted, will help society decide whether to ignore, reduce, or eliminate specific risks of genetically engineered organisms.

Review

- What is the potential problem with transgenic salmon? How are scientists addressing this concern?

SUMMARY WITH KEY TERMS

Learning Objectives

1 Explain how a typical restriction enzyme cuts DNA molecules, and give examples of the ways in which these enzymes are used in recombinant DNA technology (page 323).

- **Recombinant DNA technology** isolates and amplifies specific sequences of DNA by incorporating them into **vector** DNA molecules. Researchers then **clone**—propagate and amplify—the resulting recombinant DNA in organisms such as *E. coli*.
- Researchers use **restriction enzymes** to cut DNA into specific fragments. Each type of restriction enzyme recognizes and cuts DNA at a highly specific base sequence. Many restriction enzymes cleave DNA sequences to produce complementary, single-stranded sticky ends.

ThomsonNOW™ Learn more about restriction enzymes by clicking on the figure in ThomsonNOW.

- Geneticists may use recombinant DNA vectors from naturally occurring circular bacteria DNA molecules called **plasmids** or from bacterial viruses called **bacteriophages.**
- Geneticists often construct recombinant DNA molecules by allowing the ends of a DNA fragment and a vector, both cut with the same restriction enzyme, to associate by complementary base pairing. Then **DNA ligase** covalently links the DNA strands to form a stable recombinant molecule.

ThomsonNOW™ Learn more about the formation of recombinant DNA by clicking on the figure in ThomsonNOW.

2 Distinguish among a genomic DNA library, a chromosome library, and a complementary DNA (cDNA) library; explain why one would clone the same eukaryotic gene from both a genomic DNA library and a cDNA library (page 323).

- A **genomic DNA library** contains thousands of DNA fragments that represent the total DNA of an organism. A **chromosome library** contains all the DNA fragments from a specific chromosome. Each DNA fragment of a genomic DNA or chromosome library is stored in a specific bacterial strain. Analyzing DNA fragments in genomic DNA and chromosome libraries yields useful information about genes and their encoded proteins.
- A **cDNA library** is produced using **reverse transcriptase** to make DNA copies of mature mRNA isolated from eukaryotic cells. These copies, known as **complementary DNA (cDNA),** are then incorporated into recombinant DNA vectors.

ThomsonNOW™ Learn more about the formation of cDNA by clicking on the figure in ThomsonNOW.

- Genes present in genomic DNA and chromosome libraries from eukaryotes contain **introns,** regions that do not code for protein. Those genes can be amplified in bacteria, but the protein is not properly expressed. Because the introns have been removed from mRNA molecules, eukaryotic genes in cDNA libraries can be expressed in bacteria to produce functional protein products.

3 Describe the purpose of a genetic probe (page 323).

- Researchers use a radioactive DNA or RNA sequence as a **genetic probe** to screen thousands of recombinant DNA molecules in bacterial cells to find the colony that contains the DNA of interest.

ThomsonNOW™ Watch the process for using a genetic probe to find bacterial cells with a specific recombinant DNA molecule by clicking on the figure in ThomsonNOW.

4 Describe how the polymerase chain reaction amplifies DNA in vitro (page 323).
- The **polymerase chain reaction (PCR)** is a widely used, automated, in vitro technique in which researchers target a particular DNA sequence by specific primers and then clone it using a heat-resistant DNA polymerase.
- Using PCR, scientists amplify and analyze tiny DNA samples taken from various sites, from crime scenes to archaeological remains.

ThomsonNOW™ Watch the amplification of DNA by PCR by clicking on the figure in ThomsonNOW.

5 Distinguish among DNA, RNA, and protein blotting (page 330).
- A **Southern blot** detects DNA fragments by separating them using gel electrophoresis and then transferring them to a nitrocellulose or nylon membrane. A probe is then **hybridized** by complementary base pairing to the DNA bound to the membrane, and the band or bands of DNA are identified by autoradiography or chemical luminescence.
- When RNA molecules that are separated by electrophoresis are transferred to a membrane, the result is a **Northern blot.**
- A **Western blot** consists of proteins or polypeptides previously separated by gel electrophoresis.

6 Describe the chain termination method of DNA sequencing (page 330).
- **DNA sequencing** yields information about the structure of a gene and the probable amino acid sequence of its encoded proteins. Geneticists compare DNA sequences with other sequences stored in massive databases.
- Automated DNA sequencing is based on the **chain termination method,** which uses dideoxynucleotides, each tagged with a different-colored fluorescent dye, to terminate elongation during DNA replication. Gel electrophoresis separates the resulting fragments, and a laser identifies the nucleotide sequence.

ThomsonNOW™ Learn more about automated DNA sequencing by clicking on the figure in ThomsonNOW.

7 Describe the three main areas of interest in genomics (page 333).
- **Genomics** is the emerging field of biology that studies the entire DNA sequence of an organism's **genome.** Structural genomics is concerned with mapping and sequencing genomes. Functional genomics deals with the functions of genes and nongene sequences in genomes. Comparative genomics involves comparing the genomes of different species to further our understanding of evolutionary relationships.

8 Explain what a DNA microarray does, and give an example of its research and medical potential (page 333).
- Many diagnostic tests involve **DNA microarrays,** in which hundreds of different DNA molecules are placed on a glass slide or chip. DNA microarrays enable researchers to compare the activities of thousands of genes in normal and diseased cells from tissue samples. Because cancer and other diseases exhibit altered patterns of gene expression, DNA microarrays have the potential to identify disease-causing genes or the proteins they code for, which can then be targeted by therapeutic drugs.

9 Define *pharmacogenetics* and *proteomics* (page 333).
- **Pharmacogenetics,** the new science of gene-based medicine, takes into account the subtle genetic differences among individuals and customizes drugs to match a patient's genetic makeup.
- The study of all the proteins encoded by the human genome and produced in a person's cells and tissues is **proteomics.** Scientists want to identify all the proteins made by a given kind of cell, but the process is much more complicated than sequencing the human genome.

10 Describe at least one important application of recombinant DNA technology in each of the following fields: medicine and pharmacology, DNA fingerprinting, and transgenic organisms (page 338).
- Genetically altered bacteria produce many important human protein products, including insulin, growth hormone, tissue plasminogen activator (TPA), tissue growth factor-beta (TGF-β), clotting factor VIII, and Dornase Alpha (DNase).
- **DNA fingerprinting** is the analysis of an individual's DNA. It is based on a variety of **short tandem repeats (STRs),** molecular markers that are highly **polymorphic** within the human population. DNA fingerprinting has applications in such areas as law enforcement, issues of disputed parentage, and tracking tainted foods.

ThomsonNOW™ Solve a murder case with DNA fingerprinting by clicking on the figure in ThomsonNOW.

- **Transgenic organisms** have foreign DNA incorporated into their genetic material. Transgenic livestock produce foreign proteins in their milk. Transgenic plants have great potential in agriculture.

ThomsonNOW™ Learn more about the development of transgenic rice by clicking on the figure in ThomsonNOW.

11 Describe at least two safety issues associated with recombinant DNA technology, and explain how these issues are being addressed (page 342).
- Some people are concerned about the safety of genetically engineered organisms. To address these concerns, scientists carry out recombinant DNA technology under specific safety guidelines.
- The introduction of transgenic plants and animals into the natural environment, where they may spread in an uncontrolled manner, is an ongoing concern that must be evaluated on a case-by-case basis.

TEST YOUR UNDERSTANDING

1. A plasmid (a) is used as a DNA vector (b) is a type of bacteriophage (c) is a type of cDNA (d) is a retrovirus (e) b and c
2. DNA molecules with complementary sticky ends associate by (a) covalent bonds (b) hydrogen bonds (c) ionic bonds (d) disulfide bonds (e) phosphodiester linkages
3. Human DNA and a particular plasmid both have sites that are cut by the restriction enzymes *Hin*dIII and *Eco*RI. To make recombinant DNA, the scientist should (a) cut the plasmid with *Eco*RI and the human DNA with *Hin*dIII (b) use *Eco*RI to cut both the plasmid and the human DNA (c) use *Hin*dIII

to cut both the plasmid and the human DNA (d) a or b (e) b or c

4. Which of the following sequences is *not* palindromic?

   (a) 5′—AAGCTT—3′
   3′—TTCGAA—5′

   (b) 5′—GATC—3′
   3′—CTAG—5′

   (c) 5′—GAATTC—3′
   3′—CTTAAG—5′

   (d) 5′—CTAA—3′
   3′—GATT—5′

   (e) b and d

5. Which technique rapidly replicates specific DNA fragments without cloning? (a) gel electrophoresis (b) cDNA libraries (c) genetic probe (d) restriction fragment length polymorphism (e) polymerase chain reaction

6. The PCR technique uses (a) heat-resistant DNA polymerase (b) reverse transcriptase (c) DNA ligase (d) restriction enzymes (e) b and c

7. A cDNA clone contains (a) introns (b) exons (c) anticodons (d) a and b (e) b and c

8. The dideoxynucleotides ddATP, ddTTP, ddGTP, and ddCTP are important in DNA sequencing because they (a) cause premature termination of a growing DNA strand (b) are used as primers (c) cause the DNA fragments that contain them to migrate more slowly through a sequencing gel (d) are not affected by high temperatures (e) have more energy than deoxynucleotides

9. In the Southern blot technique, _____ is/are transferred from a gel to a nitrocellulose or nylon membrane. (a) protein (b) RNA (c) DNA (d) bacterial colonies (e) reverse transcriptase

10. Gel electrophoresis separates nucleic acids on the basis of differences in (a) length (molecular weight) (b) charge (c) nucleotide sequence (d) relative proportions of adenine and guanine (e) relative proportions of thymine and cytosine

11. The Ti plasmid, carried by *Agrobacterium tumefaciens,* is especially useful for introducing genes into (a) bacteria (b) plants (c) animals (d) yeast (e) all eukaryotes

12. A genomic DNA library (a) represents all the DNA in a specific chromosome (b) is made using reverse transcriptase (c) is stored in a collection of recombinant bacteria (d) is a DNA copy of mature mRNAs (e) allows researchers to amplify a tiny sample of DNA

13. Tissue growth factor-beta (a) is a genetic probe for recombinant plasmids (b) is a product of DNA technology used in tissue engineering (c) is necessary to make a cDNA library (d) cannot be synthesized without a heat-resistant DNA polymerase (e) is isolated by the Southern blot technique

14. These highly polymorphic molecular markers are useful in DNA fingerprinting: (a) short tandem repeats (b) cloned DNA sequences (c) palindromic DNA sequences (d) cosmid cloning vectors (e) complementary DNAs

CRITICAL THINKING

1. What are some of the problems that might arise if you were trying to produce a eukaryotic protein in a bacterium? How might using transgenic plants or animals help solve some of these problems?
2. Would genetic engineering be possible if we did not know a great deal about the genetics of bacteria? Explain.
3. What are some of the environmental concerns regarding transgenic organisms? What kinds of information does society need to determine if these concerns are valid?
4. How is proteomics related to functional genomics?
5. **Evolution Link.** DNA technology, such as the production of transgenic animals, is only possible because widely different organisms have essentially identical genetic systems (DNA ⟶ RNA ⟶ protein). What is the evolutionary significance of identical genetic systems in organisms as diverse as bacteria and pigs?

Additional questions are available in ThomsonNOW at www.thomsonedu.com/login