GLOBAL EDITION

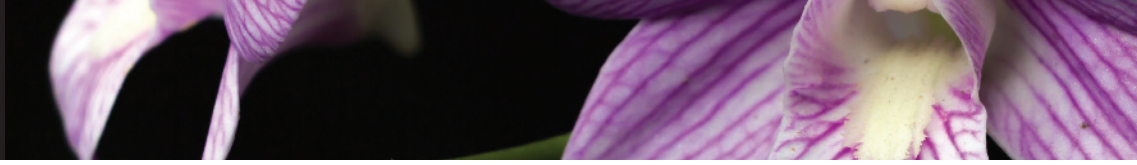
# Biology

## *A Global Approach*

TWELFTH EDITION

Campbell • Urry • Cain
Wasserman • Minorsky • Orr

# Brief Contents



CONCEPT 19.1

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

The discovery of the structure of the DNA molecule, and specifically the recognition that its two strands are complementary to each other, opened the door for the development of DNA sequencing and other techniques for manipulating DNA—known as **DNA technology—**used in biological research today. Key to these techniques is **nucleic acid hybridization**, the base pairing of one strand of a nucleic acid to a complementary sequence from another nucleic acid strand, either DNA or RNA. Nucleic acid hybridization forms the foundation of virtually every technique used in **genetic engineering**, the direct manipulation of genes for practical purposes. Genetic engineering has launched a revolution in fields as varied as criminal law, medicine, and basic biological research. In this section, we'll explore several important techniques and their uses.

## DNA Sequencing

Researchers can exploit the principle of complementary base pairing to determine the complete nucleotide sequence of a DNA molecule, a process called **DNA sequencing**. The first automated procedure, called dideoxy sequencing, was developed in the 1970s by biochemist Frederick Sanger, who received the Nobel Prize in 1980 for this accomplishment. Dideoxy sequencing is still used for routine small-scale sequencing jobs.

**Mastering Biology HHMI Video: Sanger Method of DNA Sequencing**

During the first decade of this century, "next-generation sequencing" techniques were developed that are rapid and inexpensive **(Figure 19.2)**. DNA fragments are amplified (copied) to yield an enormous number of identical fragments **(Figure 19.3)**. A single template strand of each fragment is immobilized, and the complementary strand is synthesized, one nucleotide at a time. A chemical technique enables electronic monitors to identify in real time which of the four nucleotides is added; this method is thus called *sequencing by synthesis*. Thousands or hundreds of thousands of fragments, each about 300 nucleotides long, are sequenced in parallel in machines like those shown in Figure 19.2, accounting for the high rate of nucleotides sequenced per hour. This is an example of "high-throughput" DNA technology and is currently the method of choice for studies where massive numbers of

▼ **Figure 19.2 Next-generation DNA sequencing machines.**

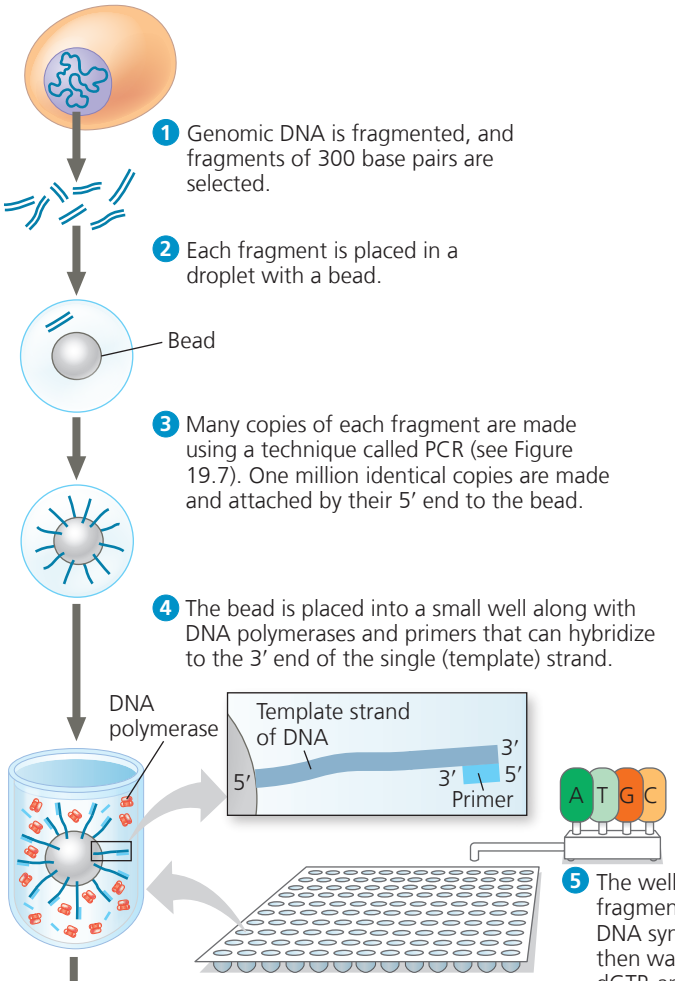
DNA samples—even a set of numerous fragments representing an entire genome—are being sequenced.

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

Improved DNA-sequencing techniques have transformed the way in which we can explore fundamental biological questions about evolution and how life works (see Make Connections Figure 5.26). Little more than 15 years after the human genome sequence was announced, researchers had completed the sequencing of thousands of genomes, with tens of thousands in progress. Complete genome sequences have been determined for cells from several cancers, for ancient humans, and for the many bacteria that live in the human intestine. In Chapter 20, you'll learn more about how this rapid acceleration of sequencing technology has revolutionized our study of the evolution of species and the genome itself. Now, let's consider how individual genes are studied.

## ▼ Figure 19.3 Research Method

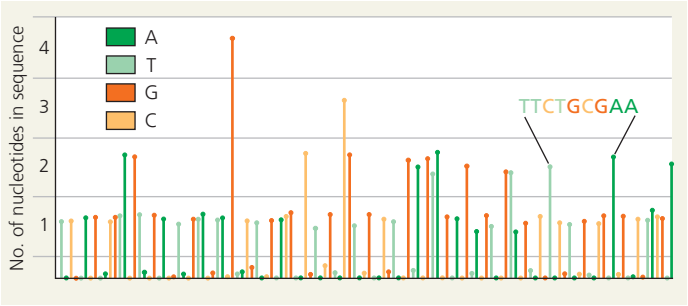
### Sequencing by Synthesis: Next-Generation Sequencing




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

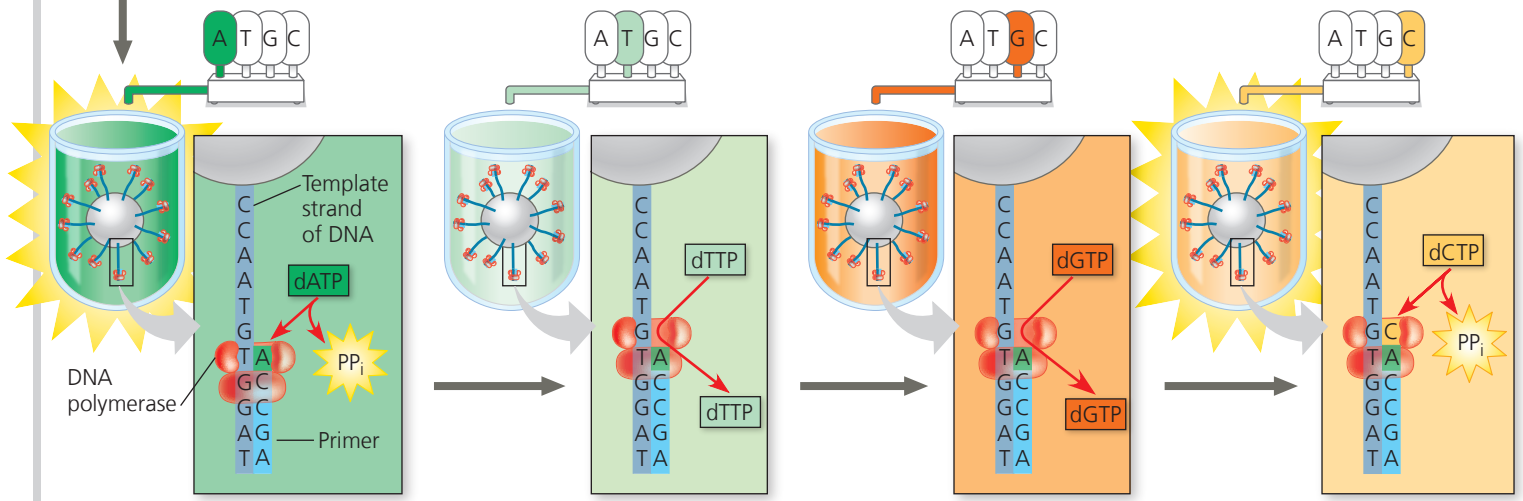
**Technique** See numbered steps and diagrams.

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




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

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




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

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

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

## Making Multiple Copies of a Gene or Other DNA Segment

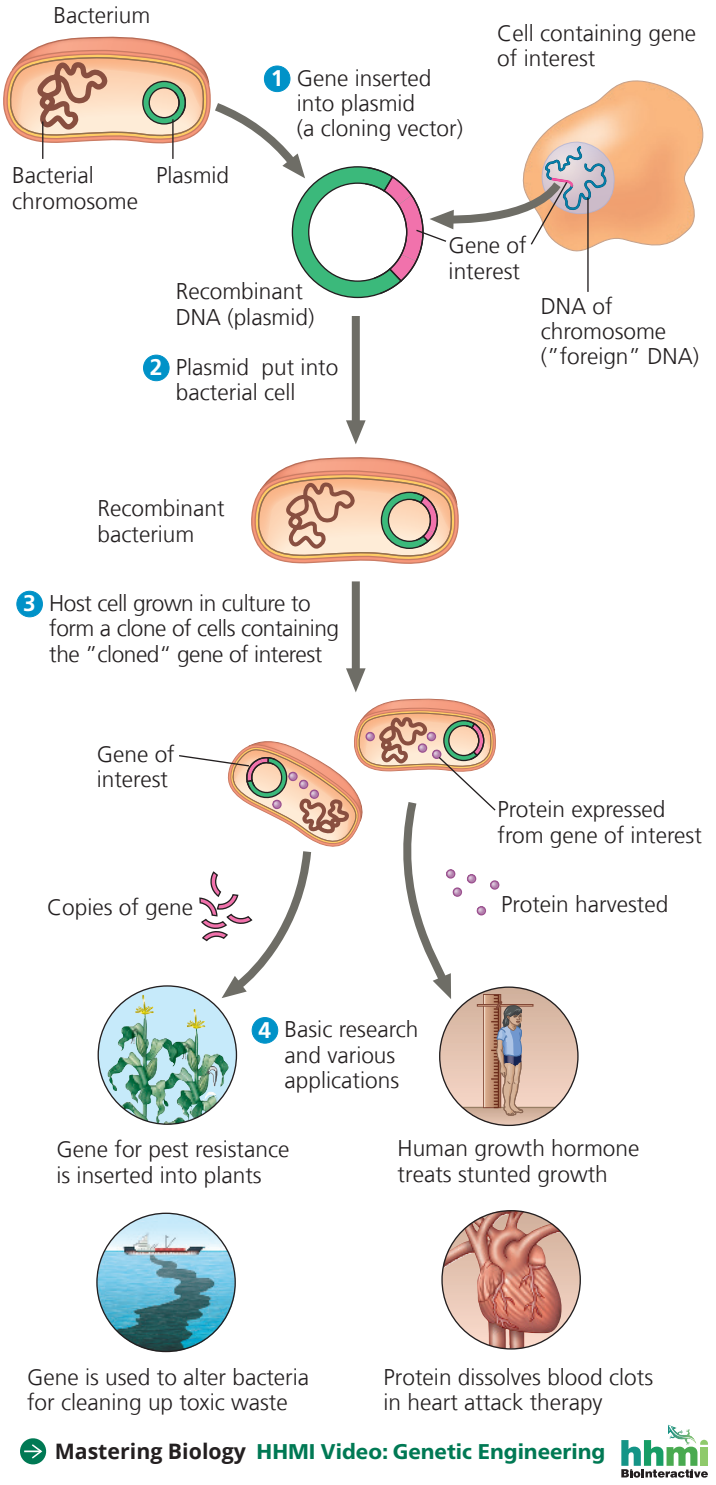
A molecular biologist studying a particular gene or group of genes faces a challenge. Naturally occurring DNA molecules are very long, and a single molecule usually carries hundreds or even thousands of genes. Moreover, in many eukaryotic genomes, protein-coding genes occupy only a small proportion of the chromosomal DNA, the rest being noncoding nucleotide sequences. A single human gene, for example, might constitute only 1/100,000 of a chromosomal DNA molecule. As a further complication, it's not easy to distinguish a gene from the surrounding DNA because they differ only in nucleotide sequence. To study a specific gene, scientists have developed methods to isolate a segment of DNA carrying that gene and make multiple identical copies of it—a process called **DNA cloning**.

Most methods for cloning pieces of DNA in the laboratory share certain general features. One common approach uses bacteria, most often *Escherichia coli*. Recall from Figure 16.13 that the *E. coli* chromosome is a large, circular molecule of DNA. In addition, *E. coli* and many other bacteria also have **plasmids**, small, circular DNA molecules that are replicated separately. A plasmid has only a small number of genes; these genes may be useful when the bacterium is in a particular environment but may not be required for survival or reproduction under most conditions.

To clone pieces of DNA using bacteria, scientists have isolated plasmids from bacterial cells and altered them by genetic engineering. Researchers insert DNA they want to study ("foreign" DNA) into the plasmid **(Figure 19.4)**. The resulting plasmid is now a **recombinant DNA molecule**, a molecule containing DNA from two different sources, very often different species. The plasmid is then returned to a bacterial cell, producing a *recombinant bacterium*. This single cell reproduces through repeated cell divisions to form a clone of cells, a population of genetically identical cells. Because the dividing bacteria replicate the recombinant plasmid and pass it on to their descendants, the foreign DNA and any genes it carries are cloned at the same time. The production of multiple copies of a single gene is a type of DNA cloning called **gene cloning**.

In Figure 19.4, the plasmid acts as a **cloning vector**, a DNA molecule that can carry foreign DNA into a host cell and be replicated there. Bacterial plasmids are widely used as cloning vectors for several reasons: They can be readily obtained from commercial suppliers, manipulated to form recombinant plasmids by insertion of foreign DNA in a test tube (referred to as *in vitro*, from the Latin meaning "in glass"), and then easily introduced into bacterial cells. The foreign DNA in Figure 19.4 is a gene from a eukaryotic cell; we will describe in more detail how the foreign DNA segment was obtained later in this section.

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




Mastering Biology **HHMI Video: Genetic Engineering** hhmi BioInteractive

Gene cloning is useful for two basic purposes: to make many copies of, or *amplify*, a particular gene and to produce a protein product from it (see Figure 19.4). Researchers can isolate copies of a cloned gene from bacteria for use in basic

research or to endow another organism with a new metabolic capability, such as pest resistance. For example, a resistance gene present in one crop species might be cloned and transferred into plants of another species. (Such organisms are called *genetically modified organisms*, or *GMOs* for short; they will be discussed later in the chapter.) Alternatively, a protein with medical uses, such as human growth hormone, can be harvested in large quantities from cultures of bacteria carrying a cloned gene for the protein. (We'll look at the techniques for expressing cloned genes later.) Since one gene is only a very small part of the total DNA in a cell, the ability to amplify such a DNA fragment is crucial for any application involving a single gene.

## Using Restriction Enzymes to Make a Recombinant DNA Plasmid

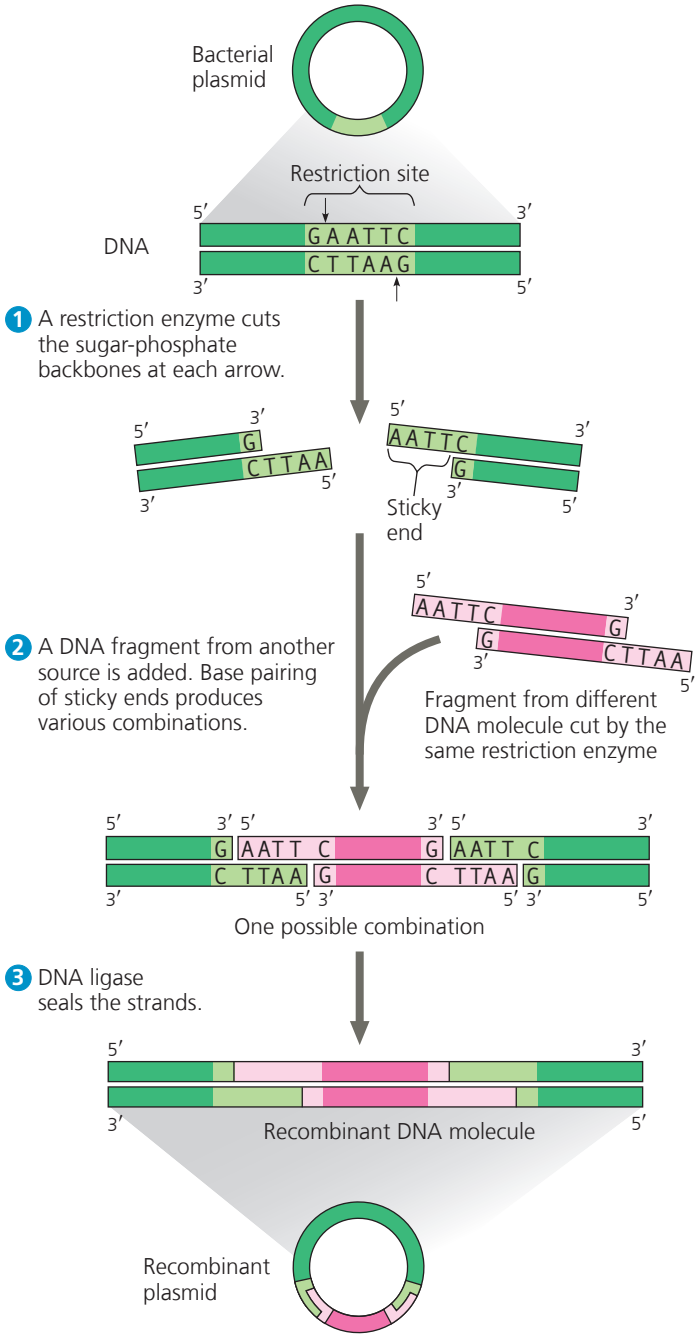
Gene cloning and genetic engineering generally rely on the use of enzymes that cut DNA molecules at a limited number of specific locations. These enzymes, called restriction endonucleases, or **restriction enzymes**, were discovered in the late 1960s by biologists doing basic research on bacteria. Restriction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

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

**Mastering Biology** Animation: Restriction Enzymes

**Figure 19.5** shows how restriction enzymes are used to clone a foreign DNA fragment into a bacterial plasmid. At the top of the figure is a bacterial plasmid (like the one shown in Figure 19.4) that has a single restriction site recognized by a particular restriction enzyme. As shown in this example, most restriction sites are symmetrical. In other words, the sequence of nucleotides is the same on both strands when read in the 5′ → 3′ direction. The most commonly used restriction enzymes recognize sequences containing four to eight nucleotide pairs. Because any sequence that is this short usually occurs (by chance) many times in a long DNA molecule, a restriction enzyme will make many cuts in such a DNA molecule, yielding a set of **restriction fragments**. Since restriction enzymes always cut at the same exact DNA sequence, copies of any given DNA molecule exposed to the same restriction enzyme always yield the same set of restriction fragments.

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




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

**Mastering Biology** Animation: Recombinant DNA

The most useful restriction enzymes cleave the sugar-phosphate backbones in the two DNA strands in a staggered manner, as shown in step 1 of Figure 19.5. The resulting double-stranded restriction fragments have at least one single-stranded end, called a **sticky end**. These short extensions can form hydrogen-bonded base pairs with complementary sticky ends on any other DNA molecules cut with the same restriction enzyme, such as the inserted DNA shown in step 2 of Figure 19.5. The associations formed in this way are only temporary but can be made permanent by DNA ligase, an enzyme that catalyzes the formation of covalent bonds that close up the sugar-phosphate backbones of DNA strands (see step 3 of Figure 19.5). At the bottom of Figure 19.5, you can see the stable recombinant DNA molecule that was produced by the ligase-catalyzed joining of DNA from two different sources. The end result, in this example, is the formation of a stable recombinant plasmid containing foreign DNA.

**Mastering Biology** Animation: Creating Recombinant DNA

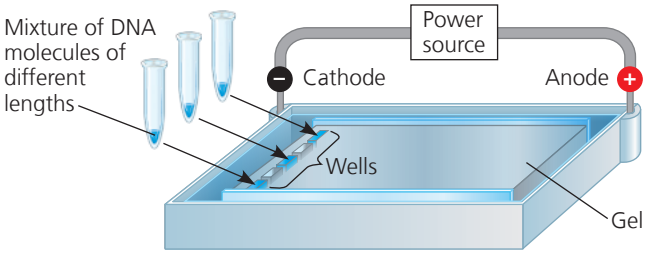
To check the recombinant plasmids after they have been copied many times in host cells to make sure the fragment has been inserted (see Figure 19.4), a researcher might cut the products again using the same restriction enzyme. If the insert is there, there would be two DNA fragments, one the size of the plasmid and one the size of the inserted DNA. To separate and visualize the fragments, researchers carry out a technique called **gel electrophoresis**, which uses a gel made of a polymer that has microscopic holes of different sizes, through which shorter fragments can travel faster. The gel works as a molecular sieve to separate out a mixture of nucleic acid fragments by length **(Figure 19.6)**. Gel electrophoresis is used in conjunction with many different techniques in molecular biology.

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

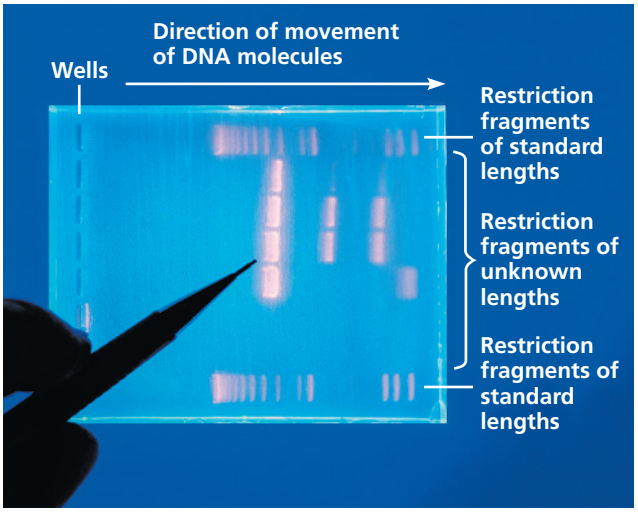
Now that we have examined the cloning vector in some detail, let's consider how biologists obtain the foreign DNA to be inserted. Most researchers have some sequence information about the DNA fragment they want to clone. Using this information, they can start with genomic DNA from the particular species of interest and obtain many copies of the desired gene by using a technique called the **polymerase chain reaction**, or **PCR**. **Figure 19.7** illustrates the steps in PCR. Within a few hours, this technique can make billions of copies of a specific target DNA segment in a sample, even if that segment makes up less than 0.001% of the total DNA in the sample.

**Mastering Biology** HHMI Video: Polymerase Chain Reaction (PCR)

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




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




**(b)** Shorter molecules are slowed down less than longer molecules, so shorter molecules move faster through the gel. After the current is turned off, a DNA-binding dye is added that fluoresces pink in ultraviolet (UV) light. Each pink band corresponds to many thousands of DNA molecules of the same length. The bands at the upper and lower edges of the gel are restriction fragments of standard lengths for comparison with samples of unknown length.

**Mastering Biology** Animation: Gel Electrophoresis of DNA

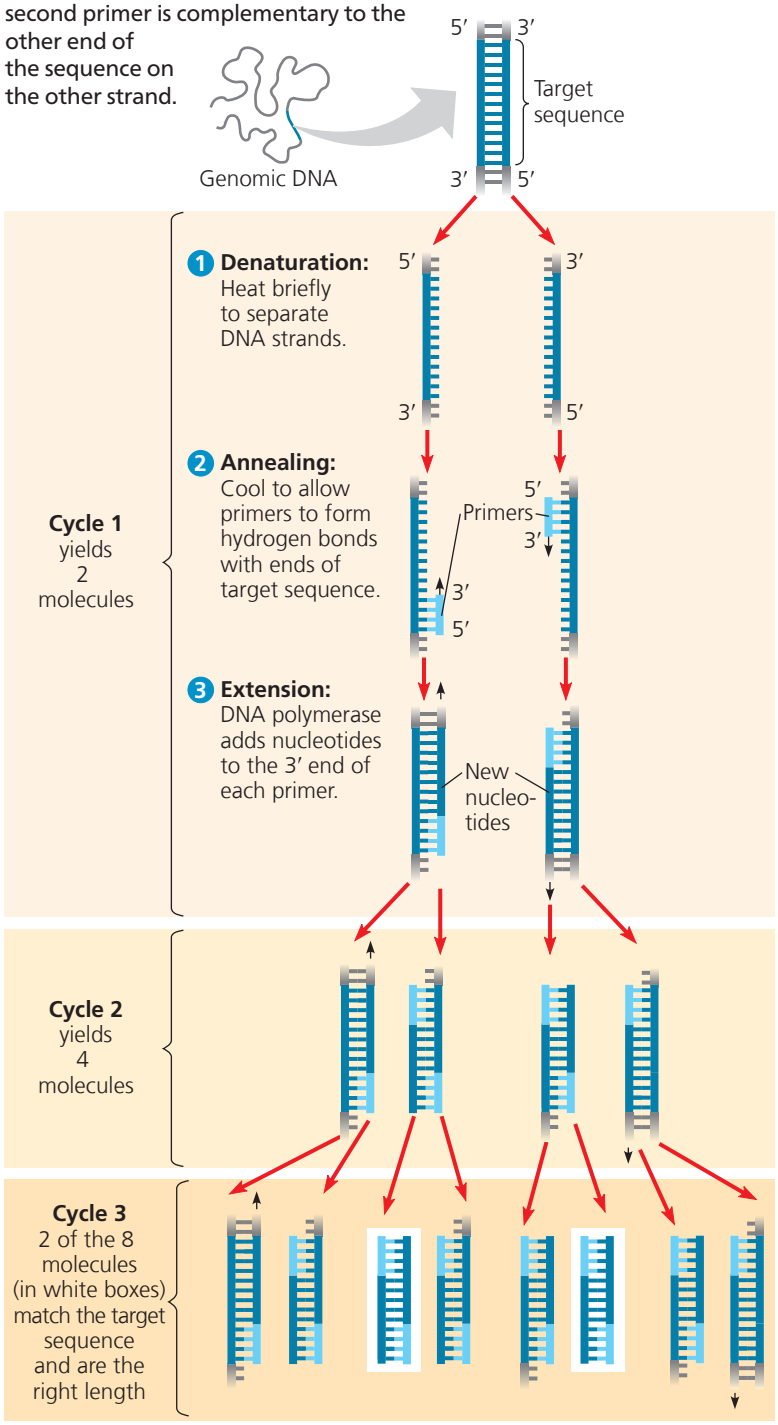
In the PCR procedure, a three-step cycle causes a chain reaction that produces an exponentially growing population of identical DNA molecules. During each cycle, the reaction mixture is 1 heated to high temperatures to denature (separate) the strands of the double-stranded DNA and then 2 cooled to allow annealing (hydrogen bonding) of short, single-stranded DNA primers complementary to sequences on opposite strands at each end of the target sequence; finally, 3 a special DNA polymerase extends the primers in the 5′ → 3′ direction. This cycle is then repeated 30–40 times. If a standard DNA polymerase were used, this enzyme

▼ **Figure 19.7 Research Method**

## The Polymerase Chain Reaction (PCR)

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

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




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

**Mastering Biology** **Animation: Copying DNA Through PCR**

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

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

Despite its speed and specificity, PCR amplification cannot substitute for gene cloning in cells to make large amounts of a gene. This is because the polymerases that are used have no proofreading function, and occasional errors during PCR replication limit the number of good copies and the length of DNA fragments that can be copied. Instead, PCR is used to provide the specific DNA fragment for cloning. PCR primers are synthesized to include a restriction site at each end of the DNA fragment that matches the site in the cloning vector, and the fragment and vector are cut and ligated together **(Figure 19.8)**. The resulting plasmids are sequenced so that those with error-free inserts can be selected.

Devised in 1985, PCR has had a major impact on biological research and genetic engineering. PCR has been used to amplify DNA from a wide variety of sources: a 40,000-year-old frozen woolly mammoth; fingerprints or tiny amounts of blood, tissue, or semen found at crime scenes; single embryonic cells for rapid prenatal diagnosis of genetic disorders (see Figure 14.19); and cells infected with viruses that are