Chapter 10

Molecular Cloning and Recombinant DNA Technology

After reading this chapter, you should be able to:

- Describe tools used to manipulate and recombine DNA
- Explain how to create novel DNA constructs

Techniques covered:

- **Isolation of DNA fragments**: restriction enzymes, polymerase chain reaction (PCR), reverse transcription PCR (RT-PCR), quantitative real-time PCR (qRT-PCR), agarose gel electrophoresis
- **Cloning DNA fragments**: vectors, ligation, transformation, DNA purification
- **Identifying DNA fragments**: DNA sequencing, Southern blot, northern blot

Molecular biology techniques allow scientists to manipulate DNA sequences in vitro. **Cloning** refers to the process of identifying and cutting a piece of DNA out of the genome and storing it for future use. **Recombinant DNA technology** refers to a suite of tools for cutting, ligating, and mass-producing DNA sequences to create useful DNA constructs for future experiments. For example, a scientist may wish to fuse the genetic sequence of a protein of interest with the gene encoding green fluorescent protein for a future experiment to visualize trafficking of the protein within cells. Using cloning and recombinant DNA techniques, the scientist can fuse the two sequences into a single DNA construct and mass produce the sequence to eventually deliver to cells.

The purpose of this chapter is to describe molecular biology tools and how they are used to clone and manipulate DNA constructs. First, we will discuss methods of isolating DNA fragments from the genome and other DNA vectors. Then, we will describe molecular cloning: how these isolated fragments are inserted into storage vectors and mass-produced using bacteria. Next, we describe how DNA fragments or recombinant DNA constructs are purified from other DNA molecules. Finally, we describe methods of reading the molecular sequence of a DNA construct. After creating and purifying a novel

Guide to Research Techniques in Neuroscience. https://doi.org/10.1016/B978-0-12-818646-6.00014-2

DNA construct, a scientist can deliver this construct into various cell types (Chapter 11) or use the construct to produce a genetically modified organism (Chapter 12).

ISOLATING DNA FRAGMENTS

In Chapter 9, we discussed methods of identifying genes that could be important for a neural phenotype. Once a gene has been identified, isolating the DNA sequence that encodes the gene is critical for subsequent experiments and techniques used to manipulate and recombine the DNA in useful ways. There are two primary methods used to isolate DNA sequences from their location in the genome or a storage vector: (1) restriction digests that cut DNA into small pieces and (2) PCR, which amplifies DNA sequences of interest.

Restriction Enzymes

The discovery of **restriction enzymes** (also called **restriction endonucleases**) that recognize and cut specific double-stranded DNA sequences is one of the most important advances in biology that makes recombinant DNA technology possible. Restriction enzymes are like targeted molecular scissors, cutting double-stranded DNA at unique base-pair sequences called **recognition** (or **restriction**) **sites**. These enzymes are naturally produced by various strains of bacteria, which normally produce these enzymes to cleave foreign DNA molecules in their environment. Molecular biologists have identified and purified hundreds of unique restriction enzymes from many species of bacteria, and these enzymes can be easily obtained from a variety of commercial manufacturers.

Each enzyme recognizes a unique sequence of about four to eight nucleotides and cuts at specific locations within these recognition sites in a process called a **restriction digest**. Some restriction enzymes digest DNA in a way that produces **blunt ends** on the newly cut strands (Fig. 10.1A). Others produce staggered cuts that leave **sticky** or **cohesive ends** with short single-stranded tails hanging over the ends of each fragment (Fig. 10.1B). Each tail can form complementary base pairs with the tail at another end produced by the same enzyme. This cutting allows any two DNA fragments to be easily joined together, as long as the fragments were generated with the same restriction enzyme (or another enzyme that leaves the same sticky ends). These different ends can be thought of as puzzle pieces whose shape is determined by the particular enzyme used. Therefore, if the same enzyme is used to cut two different pieces of DNA, then the resulting pieces will fit with each other because they have the same corresponding sticky ends that complement each other.

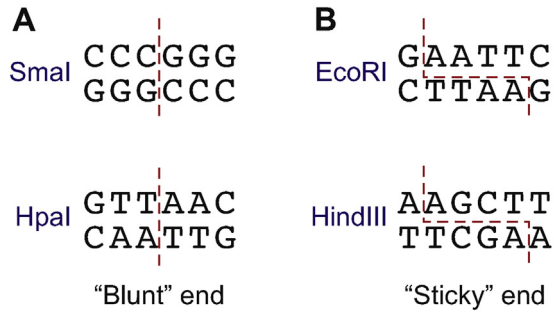

FIGURE 10.1 **Restriction enzymes.** Restriction enzymes recognize short, specific sequences of DNA and cut at precise locations within those sequences. (A) The enzymes SmaI and HpaI cut specific six base-pair sequences that result in "blunt ends" in which all bases form a pair. (B) The enzymes EcoRI and HindIII produce uneven cuts that result in "sticky ends" with overhangs. These cuts are considered sticky because the DNA can be more easily ligated with DNAs that have the complementary sticky overhangs.

Polymerase Chain Reaction

The **polymerase chain reaction (PCR)** is a biochemical reaction that uses controlled heating and cooling in the presence of DNA synthesizing enzymes to exponentially amplify a small DNA fragment. It requires only tiny amounts of DNA to be used as a template from which a selected region will be amplified a billion-fold, effectively purifying this DNA fragment away from the rest of the template DNA.

Short single-stranded oligonucleotides (about 20–30 nucleotides in length) called **primers** are essential ingredients in PCR. Primers specify what fragment will be amplified by binding to the template in the region flanking the sequence of interest (Fig. 10.2). Where do these primers come from? They are synthesized in vitro using what is called the "solid-phase" method. In this method, short DNA strands are synthesized by adding activated monomers to a growing DNA chain linked to an insoluble support. DNA chains of as many as 100 nucleotides can be synthesized using this automated method. This type of chemical DNA synthesis can also be used to create labeled oligonucleotides for nucleic acid hybridization techniques. Labs do not often chemically synthesize DNA primers themselves but rather order them from a company or core facility at their institution that specializes in oligonucleotide synthesis.

Standard PCR

Standard PCR starts with a DNA template, often from an extraced DNA sample or storage vector, and uses primers to selectively amplify a specific DNA fragment of interest. The ingredients of a PCR reaction include a DNA template containing the target sequence to be amplified; a pair of

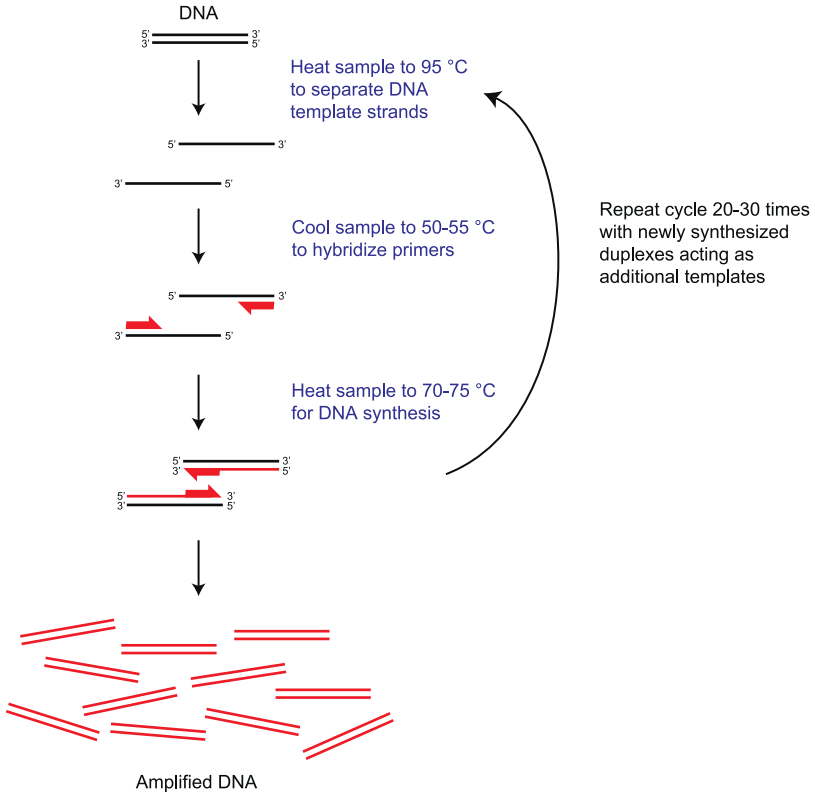

FIGURE 10.2 **Polymerase chain reaction (PCR).** Double-stranded DNA is heated to 95°C to separate the strands. Next, the temperature is lowered to about 50–55°C to allow primers to hybridize to their complementary sequences on the separated DNA strands. The temperature is raised to about 70–75°C so that Taq DNA polymerase can add nucleotides to the growing 3′ end of the primer. This cycle is repeated 20–30 times until the DNA sequence is exponentially amplified.

oligonucleotide primers that hybridize with the sequences flanking the target sequence; all four deoxyribonucleoside triphosphates (dNTPs—A, T, G, C); and a heat-stable DNA polymerase, an enzyme that adds nucleotides onto the end of a forming DNA strand.

A PCR cycle consists of three steps (Fig. 10.2):

1. **Strand separation**. The two strands of the template DNA molecule are separated by heating the solution to 95°C for about 15 s so that each strand can serve as a template.

2. **Hybridization of primers**. The solution is abruptly cooled to about 50–55°C to allow each primer to hybridize to a DNA strand. One primer hybridizes to the 3′-end of the target on one strand, and the other primer

hybridizes to the 3′-end on the complementary target strand so that DNA extension from the primers will be directed toward each other. This amplifies the region between the primers. Parent DNA duplexes do not form, because the primers are present in large excess.

3. **DNA synthesis**. The solution is heated to about 70−75°C, the optimal temperature for *Taq* DNA polymerase. This heat-stable polymerase is derived from *Thermus aquaticus*, a thermophilic bacterium found naturally in hot springs. The polymerase elongates both primers in the direction of the target sequence because DNA synthesis is in the 5′ to 3′ direction; new dNTPs are added to the growing 3′ end of each primer. DNA synthesis takes place on both strands and extends so that the target DNA sequence is fully extended to the length between each of the two primers.

These three steps (strand separation, hybridization of primers, and DNA synthesis) constitute one cycle of PCR amplification. A standard PCR protocol consists of 25−30 cycles, and each cycle effectively doubles the amount of the target DNA from the previous cycle. A single cycle requires only about 2−3 min (depending on the length of the PCR template to be amplified), and the entire procedure can be easily automated using modern PCR machines (known as **thermocyclers**). The end result of a PCR reaction is a solution in which the predominant DNA species are several million copies of the target DNA sequence dictated by the position of the two primers.

Reverse Transcription PCR

Reverse transcription PCR (RT-PCR) uses mRNA rather than DNA as the starting template. First, the enzyme reverse transcriptase uses the mRNA template to produce a complementary single-stranded DNA strand called **complementary DNA (cDNA)**. Next, DNA polymerase is used to convert the single-stranded cDNA into double-stranded DNA. These DNA molecules can then be used as templates for a PCR reaction as described above. The utility of RT-PCR is that it can be used to determine if an mRNA species is present in a sample or to clone a cDNA sequence for a subsequent experiment.

Quantitative Real-Time PCR

To compare the relative amount of mRNA or DNA present in different samples, a scientist can employ a modified version of PCR called **quantitative real-time PCR (qRT-PCR)**, not to be confused with RT-PCR described above. This technique is based on detecting a fluorescent signal produced in proportion to the PCR product during each cycle of DNA amplification. There are multiple methods of producing a fluorescent signal during a qRT-PCR reaction. In a TaqMan assay (Fig. 10.3), a fluorogenic probe within the

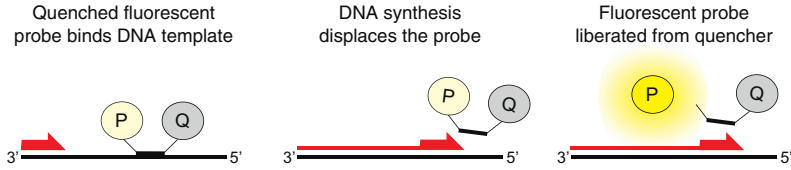

FIGURE 10.3 Quantitative real-time PCR (qRT-PCR) using a TaqMan assay. A probe containing a fluorescent probe (P) fused to a quencher (Q) hybridizes with a DNA sequence. During a PCR reaction, DNA synthesis of the growing strand cleaves the probe, causing removal of the quencher and increasing the fluorescent signal. Each round of amplification produces fluorescence that can be quantified to determine the amount of nucleic acid template present in the sample.

region to be amplified contains a fluorescent reporter tag on one end and a fluorescence quencher on the other end. During each cycle of amplification, DNA polymerase cleaves the fluorescent probe molecule, removing the quencher and allowing the reporter to emit fluorescence. The increase in fluorescence is proportional to the amplification during the exponential phase of amplification. An alternate method uses a fluorescent dye called SYBR Green that binds to double-stranded DNA; the resulting DNA—dye complex is capable of absorbing blue light and emitting green light, indicating the concentration of amplified double-stranded DNA. A sensitive light meter within the PCR chamber can measure and record the fluorescent signal after each PCR cycle and produce a real-time measure of the concentration of double-stranded amplified DNA fragments. Thus, a scientist can compare the relative concentration between different samples.

Nucleic acid hybridization techniques such as Southern (DNA) or northern (RNA) blots (discussed later in the chapter) are alternative methods to quantitatively compare amounts of DNA or RNA present in samples. However, these methods are not as quantitative as qRT-PCR and typically require large concentrations of starting material.

Modifying DNA Sequences Using PCR

Though the primary purpose of PCR is to make exact copies of a DNA sequence, this technique can also be used to modify a DNA sequence for subsequent experiments. For example, a scientist may want to introduce a point mutation into a sequence of DNA, or add base pairs to a sequence to introduce restriction sites. To modify DNA sequences using PCR, scientists design primers that not only specify the DNA fragment to be amplified, but also add or alter base pairs within the primer region itself. Though the primers may not hybridize with the template DNA strands as tightly for the first few PCR cycles, with each successive cycle, the added or mutated sequences in the primers will be copied and incorporated into the amplified PCR product.

Isolation and Characterization of DNA Fragments Using Gel Electrophoresis

Once DNA fragments have been generated using restriction enzymes or PCR amplification, they must be purified and separated from other DNA fragments that are not of interest. The separation and identification of DNA fragments based on their size is possible using a ubiquitous tool called gel electrophoresis. **Gel electrophoresis** is used to isolate, identify, and characterize properties of DNA fragments (Fig. 10.4).

First, a scientist uses molds to make an agarose gel with wells at one end for placing DNA samples. The gel is placed in an electrophoresis chamber and DNA samples are added to the wells. The electrophoresis apparatus produces a small electrical field, driving negatively charged DNA strands away from the cathode (the negative end) and toward the anode (the positive end). The mobility of the DNA fragments through the gel is inversely proportional to the logarithm of the number of base pairs in the fragments. Therefore, smaller DNA fragments will migrate much more quickly through the gel than large fragments. One of the lanes of the gel typically contains a **DNA ladder**—DNA that has been previously digested and characterized, producing fragments with

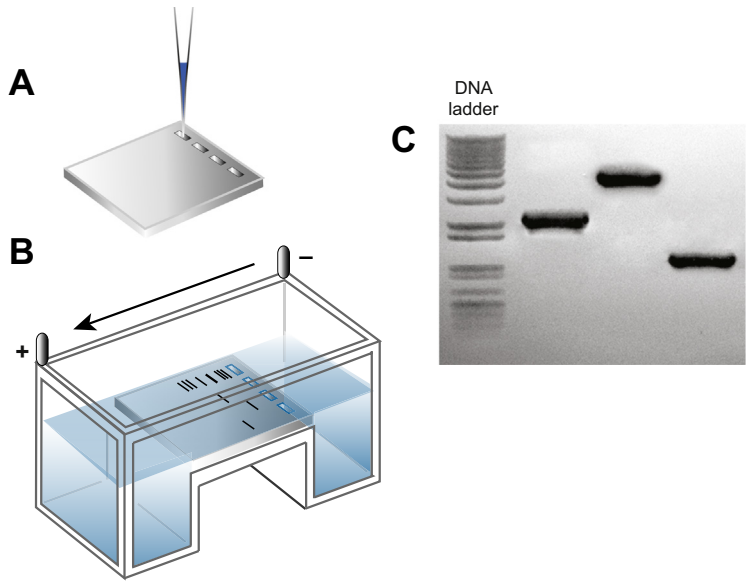

FIGURE 10.4 **Gel electrophoresis.** (A) DNA samples are loaded into individual wells of an agarose gel. (B) When an electric current is passed through the gel, negatively charged DNA travels away from the cathode (−) and toward the anode (+) in a size-dependent manner. (C) The gel is soaked in a solution to visualize the DNA bands. A DNA ladder of DNA fragments of known sizes is usually run adjacent to the samples to identify their sizes.

known base-pair sizes. Therefore, it is possible to determine the sizes of the fragments in the DNA samples by comparing their location to the location of the fragments in the ladder.

The DNA bands on agarose gels are invisible unless the DNA is labeled or stained. A common method of staining agarose gels is to briefly incubate them in a solution containing **ethidium bromide**, a DNA intercalating agent that fluoresces under ultraviolet light when bound to DNA. After identifying DNA fragments in a gel, it is possible to physically cut the DNA fragments out of the gel and purify the DNA from the agarose using simple, commercially available kits and protocols. After purifying the DNA, a scientist can clone the fragment into a storage vector.

CLONING DNA

Once a DNA fragment is isolated using restriction enzymes or PCR techniques, it is necessary to "store" the fragment into a vector for future use and amplification. The process of isolating a piece of DNA, ligating it into a storage vector, and producing enough copies of the vector for subsequent experiments is described below.

Vectors

A **DNA vector** is a nucleic acid storage system that can hold an isolated DNA sequence of interest. The essential features of a vector are that it can replicate autonomously in a host species, usually bacteria, and that it can be combined with other pieces of DNA. Often, the inserted DNA fragment comes from a different organism than the vector DNA, and the fusion is called a recombinant DNA construct. Two commonly used vectors are **plasmids**, naturally occurring circles of DNA that act as independently replicating accessory chromosomes in bacteria, and **bacteriophage** or **phage**, a virus that can deliver its genetic cargo to bacteria. It is much more common for scientists performing recombinant DNA techniques in the laboratory to use plasmids over phage vectors. Many plasmids have been ingeniously modified to enhance the delivery of recombinant DNA molecules into bacteria and to facilitate the selection of bacteria harboring these vectors.

Because there is a size limit to the number of base pairs that can be stably inserted into a plasmid or phage vector (typically 1−25 kb), scientists sometimes utilize other vectors that can hold larger amounts of DNA. A cosmid is a plasmid that contains phage sequences that allow the vector to be packaged and transmitted to bacteria like a phage vector. Cosmids are more stable than regular plasmids, and therefore they can hold larger inserts (30−50 kb). Artificial chromosomes containing telomeric and centromeric regions can hold even larger inserts.

Common Vector Features

Useful vectors contain three main features:

- **An origin of replication**. An origin of replication is a specific DNA sequence at which DNA replication is initiated. The origin is therefore absolutely essential for the amplification of the vector inside a bacterial host.

- **Useful restriction enzyme sites**. Restriction sites are needed to place a DNA sequence of interest into a vector. Because vectors are designed to be easy storage systems for cloned DNA, a standard feature in many vectors is the presence of a **multiple cloning site**, a region with multiple useful restriction enzyme sites to make a compatible digest of the vector and DNA fragment more feasible. Both the vector and the DNA fragment are cut using the same restriction enzymes so that sticky ends from the DNA fragment match the sticky ends on the vector (Fig. 10.5). Then, the two molecules can be stitched together through DNA ligation. The recombinant DNA molecule comprising the vector and the DNA insert is also known as a **DNA construct.**

- **A selectable marker**. A selectable marker is a method of allowing bacteria containing the vector to be readily identified and purified. Plasmids often confer antibiotic resistance to the bacteria, so only bacteria containing the vector will survive treatment with antibiotics. Common antibiotics used in the laboratory include ampicillin and kanamycin, and many plasmids contain genes that confer ampicillin and kanamycin resistance.

Types of Vectors

Cloning vectors provide a backbone for the DNA insert to be reproduced and propagated in bacteria; however, these vectors are only useful for storing a genetic sequence. By themselves, they are incapable of allowing for transcription and translation of the gene into a functional protein product.

For a gene to give rise to a protein product, an **expression vector** must be used that contains the necessary elements for a host cell to transcribe and translate the gene. In the case of a mammalian cell, a standard mammalian expression vector will contain an origin of replication, multiple cloning site, and selectable marker, as described above. However, the expression vector will also need a promoter found in mammalian cells that can drive the expression of the gene. The coding DNA needs other features to be transcribed and translated, such as the polyadenylation tail that normally appears at the end of transcribed pre-mRNA and a sequence that attracts the ribosome for translation.

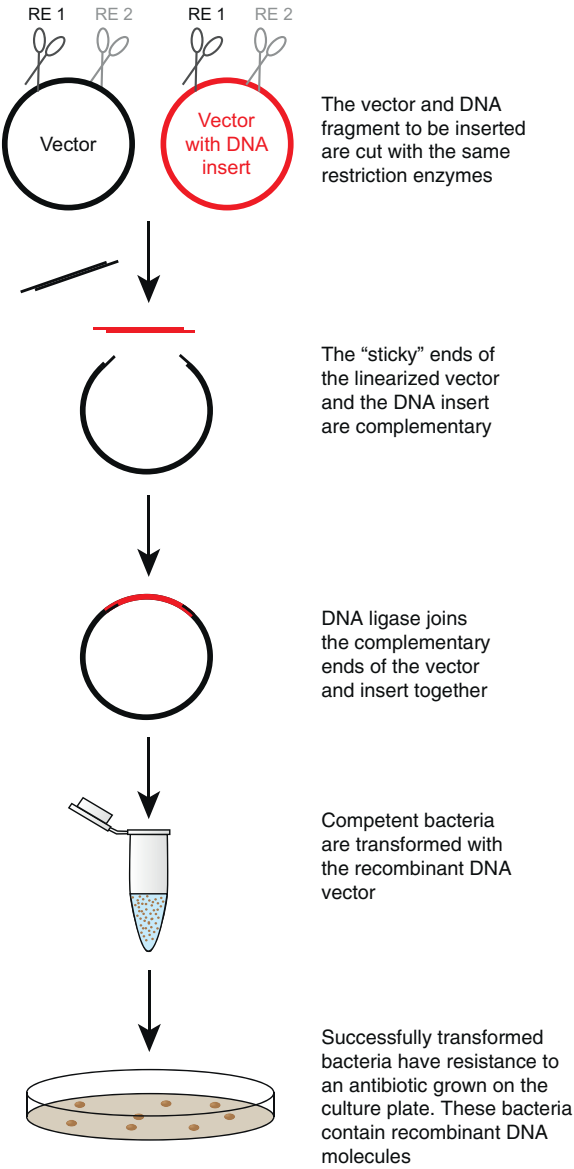

FIGURE 10.5 **Steps in making a recombinant DNA plasmid.** In this example, a scientist wishes to cut out a DNA sequence of interest from its storage vector (red) and place it into a separate vector (black). The same two restriction enzymes (REs) are used to cut each vector to produce the same "sticky end" cuts. This results in open vectors and DNA fragments that have complementary overhangs. DNA ligase is used to join the complementary ends of the vector and DNA sequence. This vector is introduced into competent bacteria and placed on an agar plate. The plate has an antibiotic that kills bacteria which do not contain the vector. Bacteria that have successfully been transformed with the vector are able to survive because the vector contains a gene that confers antibiotic resistance. The scientist can then collect bacteria colonies that contain the recombinant DNA sequence, culture large quantities of bacteria, and then lyse the bacteria to collect and purify the DNA.

Ligation

To place a DNA fragment into a vector, both the fragment and vector are usually cut using the same restriction enzymes so that they have the same sticky end fragments. The fragment can then be "pasted" into the vector using DNA ligase, an enzyme normally used by cells for DNA repair (Fig. 10.5). It is also possible to ligate DNA fragments and vectors together from restriction enzyme reactions that produce blunt end fragments, but because the blunt ends can join in multiple orientations and configurations, there is a high chance of ligation in the wrong orientation, or even the potential for multiple DNA fragments to insert into the same vector backbone.

Transformation

After a DNA fragment and vector are ligated together, the construct is introduced into **competent cells**, host cells (usually bacteria) that have been made transiently permeable to DNA (Fig. 10.5). The process of introducing a DNA vector into competent cells is called **transformation**. Bacteria can be used to grow large quantities of a single recombinant DNA molecule because they divide and grow exponentially, doubling in number every 30 min. When the host cell divides, the recombinant plasmids also replicate such that a high concentration of the desired DNA will be produced.

After transformation, the bacteria are usually grown on agar plates with an antibiotic such that only bacteria successfully carrying the vector (containing an antibiotic resistance gene) survive.

Purifying DNA from Host Cells

After transforming bacterial host cells to grow many copies of DNA, the recombinant DNA plasmid must be purified from the host cell DNA and proteins. A scientist typically takes a sample of a bacterial colony from an agar plate and grows the bacteria in fresh culture media so that the bacteria continue to divide and produce many copies of the desired DNA. Depending on the volume of bacterial culture, different amounts of DNA can be purified ranging from micrograms to grams. A scientist performs a routine protocol to lyse the bacteria, extract the DNA, and separate plasmid DNA from host cell DNA. In fact, many kits are now available from commercial manufacturers that greatly simplify the process of purifying DNA from transformed bacteria. This procedure is known as the miniprep (for small amounts derived from 1 to 3 mL starting bacterial culture), maxiprep (for 200–300 mL starting culture), or gigaprep (for 1–2 L starting culture).

IDENTIFYING DNA

After a complete cloning or subcloning experiment, it is a good idea to verify that the recombinant DNA product turned out as planned, typically by

sequencing the DNA product. DNA sequencing identifies the individual base pairs that make up the DNA strand. DNA identification can also be achieved using classical nucleic acid hybridization techniques to probe whether a specific sequence is present in a sample.

DNA Sequencing

DNA sequencing is used to determine the exact sequence of nucleotides (A,G, C, T) in a strand of DNA about 600−800 bp long. The **Sanger dideoxy chain termination method** can determine the sequence of nucleotides with high fidelity for a stretch of about 200−500 base pairs in any purified DNA sample. Based on in vitro DNA synthesis, the Sanger method synthesizes short pieces of DNA in the presence of nucleotide bases to which other bases cannot be added: chain-terminating dideoxyribonucleoside triphosphates. These chain-terminating nucleotides are labeled and mixed with regular nucleotide bases so that fragments of DNA will be created at many different lengths, each randomly stopped by the addition of a chain-terminating nucleotide. The four chain-terminating nucleotides are each labeled with a different colored fluorescent dye. Thousands of fragments of different lengths are run on a gel, and an automated fluorescence detector can quickly scan the gel to read the identity of the last, terminating nucleotide (Fig. 10.6). Modern molecular biology labs tend not to run sequencing reactions themselves, as it is typically faster and less expensive to send DNA samples off to a dedicated facility for sequencing.

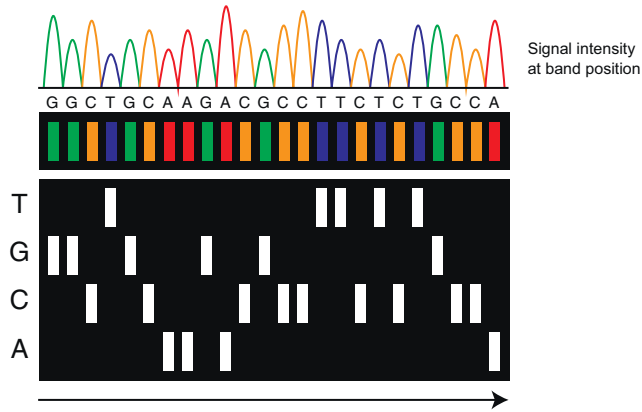

FIGURE 10.6 **DNA sequencing.** Each chain-terminating nucleotide is labeled with a different fluorescent dye, and many fragments of DNA are synthesized. When all of the synthesized DNA fragments are separated using gel electrophoresis, they will be separated by size and a detector can automatically detect which label is at each position by measuring the intensity of each fluorescent signal.

The Sanger method is the most common method of DNA sequencing for routine, day-to-day verification of molecular biology experiments. However, newer, more sophisticated **high-throughput sequencing** (also called **next-generation sequencing**) techniques can efficiently sequence very long sequences of DNA, including entire genomes (Chapter 9). These techniques are useful in the field of **genomics**, the study of the structure, function, evolution, and mapping of genomes. They are also useful for gene expression studies using **RNA-seq**, the identification of all RNA molecules in a sample.

Nucleic Acid Hybridization Techniques

The term *hybridization* refers to the phenomenon whereby two complementary nucleic acid strands bind to each other. Any fragment of DNA with a known base-pair sequence can be detected by labeling a probe strand with the complementary sequence with a radioactive label or biochemical tag that can produce a colorimetric or fluorescent signal.

Southern Blot

Developed by and named after Edward Southern, a **Southern blot** is used to identify if a known DNA sequence is present in a sample by using a specific probe to identify the sequence in a large mixture of other DNA fragments. A DNA sample (typically genomic DNA) is digested with restriction enzymes, and the mixture of fragments is separated by gel electrophoresis. The DNA is denatured to form single-stranded DNA, and the DNA is transferred to a nitrocellulose membrane (Fig. 10.7). The positions of the DNA fragments in the gel are preserved on the nitrocellulose membrane so a scientist can identify fragments based on their size. The membrane is exposed to a radioactive ^{32}P-labeled single-stranded DNA probe corresponding to the sequence of interest. This probe hybridizes with the restriction fragment containing the complementary sequence, and autoradiography reveals the position of the restriction fragment-probe duplex. Therefore, a specific fragment in a mixture of millions of other fragments can be identified on the membrane.

Southern blots were traditionally used to validate the successful introduction or removal of DNA sequences in the genome of genetically modified animals. Modern PCR techniques are more sensitive and a faster alternative to identifying a DNA fragment, but these techniques can be used in conjunction to confirm a result.

Northern Blot

A technique similar to a Southern blot used to identify the presence of RNA has been humorously named the **northern blot**. An analogous technique used to detect protein expression levels using an antibody is called a **western blot** (Chapter 14). Northern blots directly compare the relative amounts of mRNA

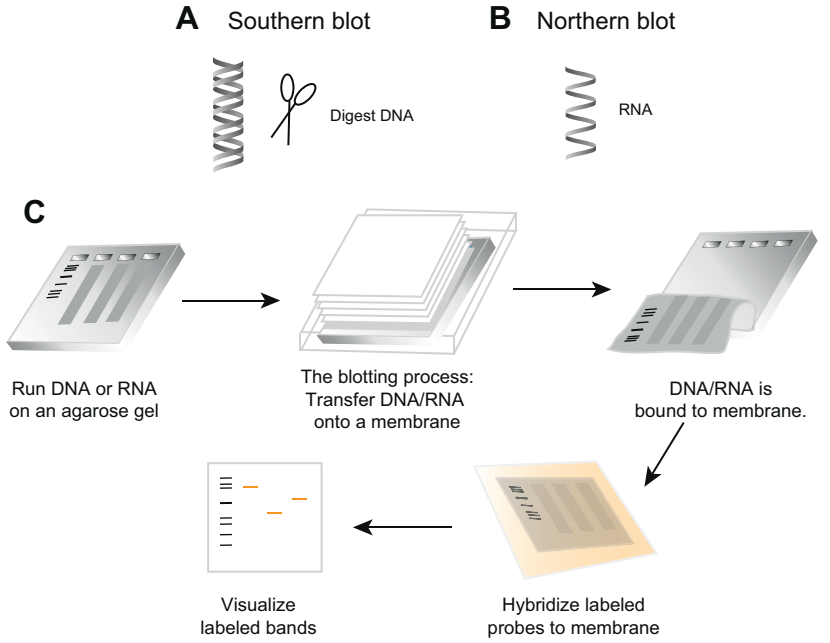

FIGURE 10.7 The Southern and northern blot methods detect specific sequences of nucleic acids. (A) In a Southern blot, restriction enzymes are used to digest DNA into fragments. (B) A northern blot uses RNA molecules as its starting material. (C) The nucleic acids are separated by size using gel electrophoresis, then transferred to a membrane. The membrane is incubated with radioactive probes capable of hybridizing to specific nucleic acid sequences. The bands can then be visualized using autoradiography.

levels between different samples. Intact mRNA molecules are separated based on their size using gel electrophoresis, and a procedure identical to that used in a Southern blot is performed to visualize the size and relative expression of the mRNA species (Fig. 10.7). Because of their efficiency and ease of use, many modern methods for examining mRNA expression levels, including qRT-PCR, in situ hybridization, and even RNA seq, have almost entirely replaced the widespread use of northern blots.

CONCLUSION

Molecular cloning and recombinant DNA technology has provided scientists with the ability to manipulate genes and the proteins they encode. These tools allow us to alter DNA sequences, cutting and pasting them in novel ways for many subsequent experiments (Box 10.1). The power of these methods is complemented by other techniques that allow for the ability to deliver DNA to cells and create transgenic and mutant organisms, as will be discussed in the next two chapters.

BOX 10.1 Walkthrough of a Subcloning Experiment

Imagine that you have identified an interesting gene called *devX* from an RNA-seq or CRISPR-Cas9 molecular screen (Chapter 9) to identify genes that decrease over the course of neural development. You identified the sequence of this gene and know that it is highly expressed in neurons. You would like to examine the subcellular trafficking of *devX* in neurons using time-lapse imaging, so you want to produce a reporter construct (Chapter 7) by tagging *devX* with green fluorescent protein (GFP) to produce *devX-GFP*. Once created, this DNA construct can be introduced into cells or animals using the transfection methods described in Chapter 11.

To create your *devX-GFP* sequence, you must first choose an appropriate vector to build your construct and determine the best restriction enzymes to use. Because you want to create a GFP fusion construct, you could use a commercially available GFP-containing plasmid vector called pEGFP-N1. This vector contains a promoter that constitutively drives gene expression (called CMV); the eGFP (enhanced green fluorescent protein) coding sequence; a gene that confers antibiotic resistance to kanamycin; and a multiple cloning site (MCS), all contained within a circular plasmid. The MCS is located in front of the *eGFP* coding sequence, meaning the *devX* coding sequence will be attached to the N-terminus of *eGFP* (this is what the N1 in the vector's name refers to).

An MCS contains many unique restriction sites to choose from, so compatible restriction enzymes can be used on both the vector and the insert. The primary factor to consider when choosing the appropriate restriction enzyme to use is whether the site is unique. You obviously do not want to cut your DNA of interest in unexpected or undesired areas. Because you will be using the same restriction enzyme to open up the vector that you will be using to generate the ends of your insert, you must consider the sequences of both items. After using software that tells you where various restriction sites are located in your DNA plasmid, you end up choosing the common enzymes EcoRI and BamHI, because *devX* does not contain those recognition sites within its coding sequence.

The next major step is to generate a *devX* fragment with the appropriate restriction sites to prepare the insert to go into the vector. There are multiple ways to generate the small fragment of DNA that represents *devX*. If you already had a vector containing the segment of *devX*, you might be able to directly cut it out of this vector to paste into a new vector. Alternatively, you could use PCR to amplify the appropriate fragment from a biological sample, adding the recognition sequences for EcoRI and BamHI to the ends of the *devX* fragment.

After choosing a vector and preparing the insert, you will need to make them compatible by cutting them with the same restriction enzymes. You need to insert the *devX* fragment in the proper direction and the same reading frame as the *eGFP* so that they form a single protein together. You have done this by choosing two different restriction enzymes (EcoRI and BamHI) that leave "sticky" ends after digestion and by making sure that the frame of the *devX* PCR product after digestion is the same as the *eGFP* coding sequence within the vector.

Continued

BOX 10.1 Walkthrough of a Subcloning Experiment—cont'd

To isolate your digested products, you run the digested vector and insert on an agarose gel. You should be able to identify the appropriate bands based on their size. Then you can physically cut out these bands and purify the DNA from the gel using simple gel purification protocols or kits.

You mix the purified digested products with DNA ligase, which will ligate the fragments together. The EcoRI digest on the 5′ end of the *devX* insert matches the EcoRI digest on the 3′ end of the plasmid, while the BamHI digest on the 3′ end of the *devX* insert matches the BamHI digest on the 5′ end of the plasmid. Thus, the *devX* sequence should be inserted in the proper direction into the plasmid.

Next you can transform competent bacteria to grow many copies of your recombinant DNA plasmid. You mix the ligation product with competent cells and then spread the cells on plates containing nutrients—to help cells grow—and kanamycin, the antibiotic that cells containing your construct should be resistant against. The next morning, you should see clonal colonies of bacteria growing on the plate, and you can select individual colonies whose member cells contain identical copies of the construct.

If you place a bit of the colony in a new tube with bacterial culture medium, you can allow the bacteria to multiply over a period of 8–12 h. Then you can purify the plasmid DNA from each of the clones using a "miniprep" kit. You lyse the bacteria, extract the DNA from the lysed bacteria, and purify the plasmid DNA from the bacterial DNA using purification columns.

The purified plasmid DNA can be examined using restriction enzyme analysis. Items that can be checked using restriction digests include the presence of the insert, whether the insert was ligated in the correct direction, and whether there is more than one insert present. These questions can be answered by choosing appropriate restriction enzymes, which will not necessarily be the ones used to join the insert and vector. Finally, you verify that the sequence in your construct matches the DNA sequence of *devX-GFP* by performing sequencing reactions (Fig. 10.6).

Now that you have produced and verified your *devX-GFP* construct, you can introduce your vector into cells using techniques that will be described in the next chapter.

SUGGESTED READING AND REFERENCES

Books

Brown, T.A., 2020. Gene Cloning and DNA Analysis: An Introduction, eighth ed. Wiley Blackwell.

Dale, J.W., von Schantz, M., Plant, N., 2011. From Genes to Genomes: Concepts and Applications of DNA Technology, third ed. Wiley.

Drlica, K., 2004. Understanding DNA and Gene Cloning: A Guide for the Curious, fourth ed. Wiley.

Nicholl, D.S.T., 2008. An Introduction to Genetic Engineering, third ed. Cambridge University Press.

Review Articles

Bras, J., Guerreiro, R., Hardy, J., 2012. Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. Nat. Rev. Neurol. 13, 453–464.

Koboldt, D.C., Steinberg, K.M., Larson, D.E., Wilson, R.K., MArdis, E.R., 2013. The next-generation sequencing revolution and its impact on genomics. Cell 155, 27–38.

Marziali, A., Akeson, M., 2001. New DNA sequencing methods. Annu. Rev. Biomed. Eng. 3, 195–223.

Primary Research Articles—Interesting Examples from the Literature

Agate, R.J., Choe, M., Arnold, A.P., 2004. Sex differences in structure and expression of the sex chromosome genes CHD1Z and CHD1W in zebra finches. Mol. Biol. Evol. 21, 384–396.

Bacon, A., Kerr, N.C., Holmes, F.E., Gaston, K., Wynick, D., 2007. Characterization of an enhancer region of the galanin gene that directs expression to the dorsal root ganglion and confers responsiveness to axotomy. J. Neurosci. 27, 6573–6580.

Brenner, S., Williams, S.R., Vermaas, E.H., Storck, T., et al., 2000. *In vitro* cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. Proc. Natl. Acad. Sci. U. S. A. 97, 1665–1670.

Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E., Heintz, N., 2003. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. Nature 425, 917–925.

Paisan-Ruiz, C., Jain, S., Evans, E.W., Gilks, W.P., et al., 2004. Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. Neuron 44, 595–600.

Valenzuela, D.M., Murphy, A.J., Frendewey, D., Gale, et al., 2003. High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. Nat. Biotechnol. 21, 652–659.

Protocols

Green, J., Sambrook, J., 2012. Molecular Cloning: A Laboratory Manual, fourth ed. Cold Spring Harbor Laboratory Press.

Websites

Addgene plasmid repository: http://www.addgene.org/.

New England BioLabs catalog: http://www.neb.com/nebecomm/tech_reference.

Primer-BLAST: primer designing tool: http://www.ncbi.nlm.nih.gov/tools/primer-blast/.