CHAPTER FIVE

PCR Theory and Practice

Since the development of PCR during the mid- to late 1980s, this method has gone on to revolutionize biology, medicine, and forensics (Palumbi 1996). Nearly three decades later, PCR continues to be a workhorse in evolutionary genetics laboratories and it is an essential tool in phylogenomics. The major reason for this success is that PCR formed a synergy with Sanger sequencing because it allowed researchers to target and sequence particular genomic loci in a far more efficient manner than earlier cloning-based practices. PCR is invaluable because it enables researchers to easily and inexpensively (in time and money) generate sufficient amounts of target DNA templates for use by DNA sequencing methods.

As briefly mentioned in Chapter 1, the Sanger sequencing workflow (Chapter 6) first requires the use of PCR to copy or "amplify" a genomic target, which is usually <2 kb in length. When a PCR is finished, billions of copies of the target locus—called *PCR products* or *amplicons*—have been synthesized. In the second step, Sanger sequencing methods are used to determine the DNA sequences encoded by the PCR amplicons. Prior to the NGS era, the PCR-Sanger sequencing workflow represented the primary means for acquiring DNA sequence data and it continues to be important today. In addition to the standard PCR method used in most Sanger sequencing-based studies, a variety of other types of PCR methods such as *hot start PCR*, *long PCR*, and *reverse transcriptase (RT)-PCR* have also been developed. We will discuss each of these methods in this chapter. In Chapter 7, we will see that PCR continues to be a highly relevant methodology in NGS workflows as well. Like Sanger sequencing, some NGS approaches use PCR amplicons as input DNA templates for sequencing. However, several PCR variants such as *limited cycle PCR*, *suppression PCR*, and *bridge PCR* are also incorporated into NGS workflows. Thus, PCR remains a critically important technique in phylogenomic data acquisition.

5.1 HISTORICAL OVERVIEW

PCR was developed into a practical laboratory technique during the 1980s by researchers at the Cetus Corporation in Berkeley, California, though a concept for the method was sketched by Kleppe et al. (1971) a decade earlier (Kornberg and Baker 1992; Palumbi 1996). The initial papers, which described the modern PCR concept, provided the first empirical case studies and showed the promise of PCR (Saiki et al. 1985; Mullis et al. 1986; Mullis and Faloona 1987). In recognition for his contributions to these accomplishments, Kary Mullis was awarded the Nobel Prize in Chemistry. However, one of the major technical problems encountered during these early PCR experiments was DNA polymerase denaturation. Exposing the *Escherichia coli* Klenow Fragment of DNA polymerase I to high temperatures (95–100°C) caused loss of activity during each replication cycle (Saiki et al. 1988). The high temperatures were needed to melt hydrogen bonds of the double helix to generate single-stranded templates for replication. In order to drive a PCR of many replication cycles, fresh DNA polymerase had to be added to the reaction tubes each cycle (Mullis et al. 1986; Palumbi 1996). To solve this problem, David Gelfand tested DNA polymerases harvested from cultures of various thermophilic

bacteria because these particular bacteria must have evolved high-temperature DNA polymerase enzymes (Brock 1997). What he discovered is that the DNA polymerase from *Thermus aquaticus*, a species of bacterium discovered in a hot springs pool in Yellowstone National Park (Brock and Freeze 1969; Figure 5.1), exhibited optimal DNA synthesis activity at 70°C and, importantly, could survive repeated exposure to the 94–95°C denaturation steps of PCR (Saiki et al. 1988; Brock 1997). This remarkable enzyme is now simply referred to as "*Taq*," which is a shortened version of the bacteria's scientific name. By using *Taq*, the specificity of PCR was also greatly increased because the higher reaction temperatures permissible with *Taq* polymerase enabled the desired reaction to proceed more efficiently. This gain in efficiency is partly explained by the decrease in production of nonspecific products, which reduces competition for DNA polymerases with target DNA (Saiki et al. 1988). *Taq* also enabled PCR to become a more rapid and automated process in that once started it could complete itself over the course of hours instead of days (Saiki et al. 1988). The discovery of *T. aquaticus* and the subsequent exploitation of its remarkable heat-tolerant DNA polymerase enzymes enabled PCR and, as we will see, DNA sequencing to achieve their potential.

Despite the PCR concept being seemingly simple to understand, beginners and even some experienced researchers nevertheless tend to regard PCR as a "black box" technique. This is because the reaction, which takes place in a small plastic tube containing only 10–50 μL of liquid, may often work as desired but fail at other times for no obvious reason. When PCR performs well it seems like an "easy" molecular method to practice, but when PCR fails it can lead a researcher to waste many hours in the laboratory with little or no results to show. Rather than merely hoping a PCR will work, the researcher should confidently *expect* it to work. If you learn the principles of PCR, then you can truly master PCR and enjoy consistent success. Part of this learning includes the importance of troubleshooting failed PCRs. Rather than viewing failed PCRs as a waste of time, it is better to learn why a given PCR failed. The rewards of such troubleshooting justify the additional effort expended.

Figure 5.1. Thermal hot springs in the vicinity of the Great Fountain Geyser in Yellowstone National Park. Brock and Freeze (1969) discovered the thermophilic bacterium *T. aquaticus* in Mushroom Pool, which is located off in the distance near the edge of the forest (the area in the foreground is closed to visitors because it is a sensitive ecological area and very dangerous). The original cultures of *T. aquaticus* used for testing *Taq* polymerase in PCR were obtained here. Photo taken by the author in August 2015.

5.2 DNA POLYMERIZATION IN LIVING CELLS VERSUS PCR

Regardless of whether DNA replication is accomplished in living cells or during *in vitro* experiments such as PCR, DNA is synthesized one strand at a time by DNA polymerase-catalyzed addition of 2′-deoxynucleoside triphosphates or what are generically referred to as "dNTPs." The four standard building blocks used to replicate genomes include: deoxyadenosine triphosphate (dATP), deoxythymidine triphosphate (dTTP), deoxyguanosine triphosphate (dGTP), and deoxycytidine triphosphate (dCTP). During DNA synthesis, bases exclusively pair with other complementary bases according to the "Watson–Crick base pair rules": guanine pairs with cytosine and thymine pairs with adenine. Although important differences exist between DNA synthesis in living systems and PCR, both forms of DNA replication generally require the same four major ingredients for proper function: (1) single-stranded RNA or DNA primers; (2) single-stranded DNA (ssDNA) template sequences; (3) dNTPs; and (4) DNA polymerase enzymes with cofactor Mg^{2+} cations.

In order to achieve a deeper understanding of PCR as well as DNA sequencing technologies, it is important to understand the details of DNA synthesis chemistry in living systems. Because the four different dNTPs represent the basic building blocks for DNA synthesis, we will begin with a look at the chemical structure of dNTPs. This remarkable molecule has several chemical components, each with a particular role during DNA synthesis. Figure 5.2 shows the basic structure of a representative dNTP—a dATP, which consists of three parts: the adenine base, a 2′-deoxyribose sugar, and a 5′-triphosphate group. The chemical groups found on one side of the base (not shown) play a critical role in the specificity of base pairing

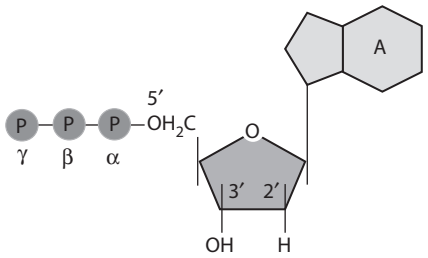

Figure 5.2. Chemical structure of a dNTP. The example shown here is dATP.

while the 3′–OH group on the sugar moiety and α-phosphate (Watson et al. 2014; Figure 5.2) are required for covalently linking together two bases on a DNA strand. The two remaining phosphates, which are labeled as γ- and β-phosphates (see Figure 5.2), interact with one of the metal ions in the active site of the DNA polymerase and later provide a source of energy to help drive the polymerization reaction. We will examine the chemistry of DNA polymerization in more detail in Section 5.2.1.

5.2.1 Brief Review of DNA Polymerization in Living Cells

Not only are many of the molecular components the same between DNA polymerization in a cell and in PCR, but the actual mechanism of polymerization is also the same. Certainly, many important differences exist between DNA synthesis in a living cell versus a man-made concoction in a plastic tube—and we will discuss these in this section. But first, let us not be distracted by these differences and instead focus on the basic polymerization reaction as it occurs in living systems because it is essentially the same as polymerization in PCR.

In a cell, after a stretch of chromosomal DNA has been separated into two single strands, DNA polymerization begins when an RNA primer binds with a ssDNA template (Figure 5.3a). Once this event occurs, a critically important *primer-template junction* is formed (Figure 5.3b). The primer-template junction enables DNA polymerase-catalyzed synthesis of DNA because it provides two critical entities: (1) a hydroxyl (-OH) chemical group (see Figure 5.2) available at the 3′ terminal nucleotide of the primer and (2) an adjacent template base. Although the correct base pairing between the template base (which is presently a thymine in Figure 5.3b) and its complement (adenine) is dictated by hydrogen bonding between the bases and their geometrical configurations relative to each other, DNA polymerase plays a critical role in helping to facilitate this pairing. After the successful base pairing between the thymine and adenine, the polymerase enzyme will then catalytically join the 3′ terminal nucleotide of the primer (a cytosine) to the adenine to complete the nucleotide addition. Note, that as each new nucleotide is added, the primer-template junction itself shifts one base along the template strand, or

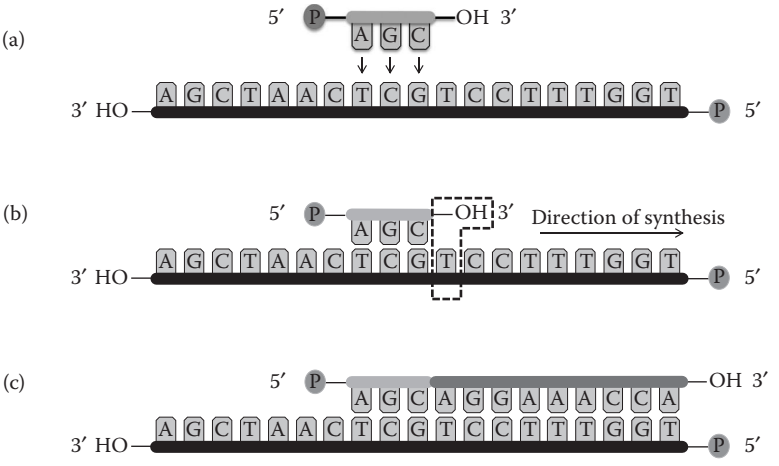

Figure 5.3. Overview of DNA synthesis. (a) A 3 bp RNA primer (light gray strand) hybridizes with the template (black) strand. (b) Primer-template junction formed (dotted box). (c) DNA polymerase synthesizes new strand (dark gray). Note, DNA polymerase and dNTPs not shown. (Modified after Figure 9-1, Page 258, of Watson, J. D. et al. 2014. *Molecular Biology of the Gene*, 7th edition. New York: Pearson Education, Inc.)

in a 5′ → 3′ direction with respect to the primer strand. Thus, each newly added nucleotide will, in turn, "prime" the addition of the next base because the 3′ end of the nontemplate strand will have the requisite 3′-OH thus allowing for further elongation of the nontemplate strand (Figure 5.3c).

Let's examine these chemical reactions in more detail. Figure 5.4 shows a 4 bp RNA primer that has just bonded to complementary nucleotides on a single-stranded template. Once the primer, a 5′–TGTC–3′ sequence in Figure 5.4, bonds to its complementary sites on the template, the DNA polymerase (not shown) will position the incoming dATP in the proper spatial orientation with respect to the template nucleotide (a thymine) and 3′-OH on the primer nucleotide (a cytosine). As soon as proper Watson–Crick base pairing occurs, the new base pair is in correct stereochemical configuration, which then allows the DNA polymerase to catalyze the reaction. In this critical step, a positively charged metal cation (Mg^{2+}) located in the active site of the polymerase enzyme removes a hydrogen atom from the 3′-OH group (Figure 5.4a). This deprotonation of the hydroxyl group converts the oxygen into a negatively charged nucleophile that then attacks the α-phosphate on the nearby dATP (Figure 5.4). This action releases a diphosphate group (β- and γ-phosphates), which is eventually split apart by the enzyme pyrophosphatase. Breaking the bond linking the two phosphates releases energy that helps drive the formation of a new covalent bond linking together the cytosine and adenine nucleotides (Figure 5.4). These inorganic diphosphate molecules are called "pyrophosphates" (PPi) because when the covalent bond linking the two inorganic phosphates together is broken, a significant amount of energy is released. Inorganic pyrophosphates are natural byproducts of DNA polymerization in both living systems and in PCR. In cells these products are broken down by the enzyme pyrophosphatase. However, in PCR these byproducts accumulate with the target PCR products.

Another difference between DNA synthesis in living cells and PCR concerns the primers. In living systems, short RNA primers, which are only 5–10 bp long, bind to many complementary places along a template sequence simply because their targets are abundantly scattered throughout any given template sequence (Watson et al. 2014). Their function is to help initiate DNA synthesis. In contrast, PCR primers are different for two reasons. First, PCR primers are made of DNA and not RNA. Secondly, PCR primers are usually at least 18 bases long. The reason for this is that, in addition to priming DNA synthesis in the same manner as RNA primers, DNA primers have the added crucial property of being *highly specific* for certain genomic locations. This specificity property of PCR primers arises because of their length and sequence. Table 5.1 shows us that as primer

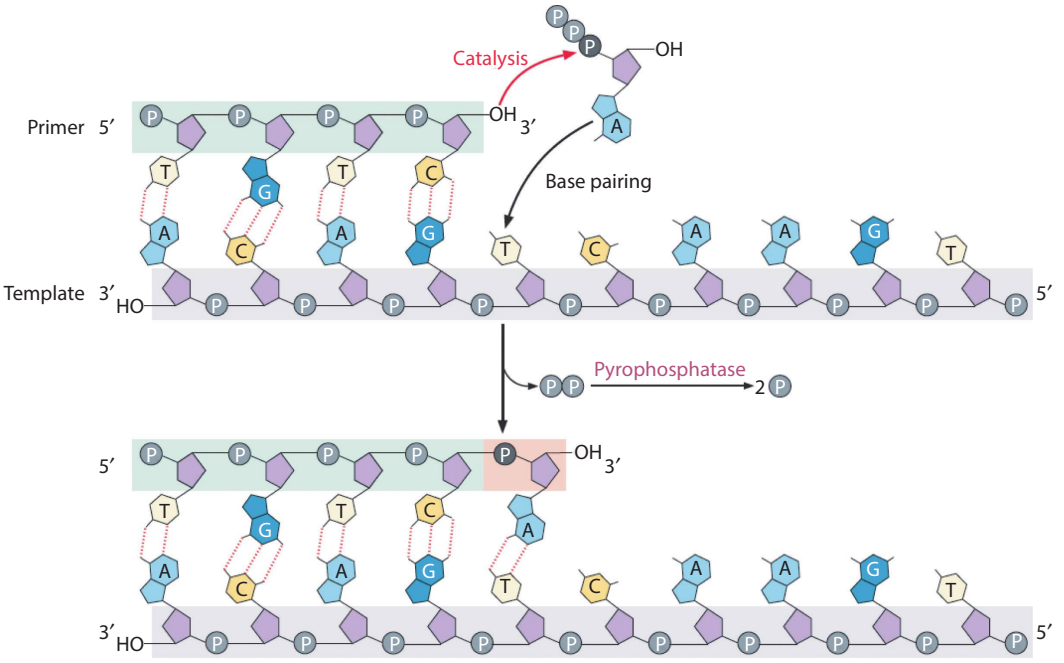

Figure 5.4. Mechanism of DNA synthesis in a living cell (Watson et al. 2014, fig. 9-2). The legend reads, "Diagram of the mechanism of DNA synthesis. DNA synthesis is initiated when the 3′-OH of the primer mediates the nucleophilic attack of the α-phosphate of the incoming dNTP. This results in the extension of the 3′ end of the primer by one nucleotide and releases one molecule of pyrophosphate. Pyrophosphatase rapidly hydrolyzes released pyrophosphate into two phosphate molecules." (Watson, J. D. et al. *Molecular Biology of the Gene*, 7th Ed., © 2014. Reprinted by permission of Pearson Education, Inc., New York, New York.)

length is increased the probability for a random match with a template declines rapidly. As primer length increases, the probabilities seem to become so small that they might lead someone to think that primers even as long as 10 bp may be sufficient for most PCR applications—after all, the probability that a primer this long would perfectly match a sequence of the same size is only 0.000000953674. However, as we saw in Chapter 2, genomes can be quite large. If we instead take the reciprocal value of this probability, assuming a random sequence of nucleotides, we can expect to find a perfect sequence match for our 10 bp primer, on average, once every 1,048,576 bp. Given that the human genome is over 3 billion bp, we could expect our 10 bp primer to anneal to roughly 2,800 different genomic locations! Perhaps, a primer this size may work fine for a genome that is on the order of millions of bases (a bacterial genome) but not for a larger genome such as ours. This is why PCR primers are usually at least 18 bases long; an 18-base primer is expected to find a random match only once every 69 *billion* bases—far larger than most eukaryotic genomes (Brown 2007)! This ability of longer DNA primers to target a single genomic location is one of the most desirable properties of PCR. As we will see in Section 5.2.2, a third major difference between primers in living systems versus PCR is that in PCR there are *two* primers of different sequences used together in a reaction to amplify a particular target sequence that usually ranges between 200 and 2,000 bases as opposed to the many primers needed to replicate an entire genome.

5.2.2 How the PCR Works

Now that our review of DNA synthesis and introduction of PCR primers is done, we can focus on the actual PCR mechanism and laboratory procedures for performing successful PCR. A PCR experiment begins when the researcher pipets the PCR reagents (template DNA, dNTPs, reaction buffer, water, and *Taq* polymerase) into 0.2 mL plastic reaction tubes. The tubes are then placed into

TABLE 5.1
Primer length and the probability of a match with a nonhomologous template sequence

Primer length (bp)	Probability of random template match	Genome size (bp) for one random template match
1	0.25	4
2	0.0625	16
3	0.015625	64
4	0.00390625	256
5	0.0009765625	1,024
6	0.000244140625	4,096
7	0.00006103515625	16,384
8	0.0000152587890625	65,536
9	0.000003814697265625	262,144
10	0.00000095367431640625	1,048,576
11	0.00000023841857910156	4,194,304
12	0.00000005960464477539	16,777,216
13	0.00000001490116119385	67,108,864
14	0.00000000372529029846	268,435,456
15	0.00000000093132257462	1,073,741,824
16	0.00000000023283064365	4,294,967,296
17	0.00000000005820766091	17,179,869,184
18	0.00000000001455191523	68,719,476,736

NOTE: As the length of a primer increases the probability that the primer will match a random segment of DNA decreases according to the function $(0.25)^X$ (where X = primer length).

a *thermocycler*, a tabletop machine used to conduct the PCR. As its name implies, a thermocycler heats and cools the reaction mixtures. Once running, the thermocycler will typically undergo 30–40 *cycles* of heating and cooling that are controlled by a computer program. Each of these cycles, in turn, is comprised of three core steps: *denaturation*, *annealing*, and *extension* (Palumbi 1996). We will now discuss these temperature steps in a PCR cycle.

Step 1: Denaturation In this initial step the genomic DNA is heated to 94–96°C, which breaks all hydrogen bonds linking together the two complementary strands of DNA. This creates a mixture of ssDNA thus making available the primer targets on the genomic DNA templates.

Step 2: Annealing Next, the reaction mixture is gradually cooled down to a specific pre-selected temperature that depends on the optimal annealing temperatures of the primers being used, which is usually between 45°C and 65°C (we will discuss these properties of primers below and in Chapter 8). Upon arriving at the correct annealing temperature, the *forward primers* hybridize to their target locations on template strands. Meanwhile, the *reverse primers* also find and bind to their targets on the complementary template strand hundreds or thousands of bases distant from the forward primer. Note that the "forward" and "reverse" primer designations usually do not have any functional significance. The researcher usually designates which is which. What is important to keep in mind is that each pair of PCR primers usually consists of one forward and one reverse primer. The specificity of primer-template matches is influenced by the primer sequences, annealing temperature, and length of time spent at the annealing temperature.

Step 3: Extension After the forward and reverse primers bind to their targets the mixture is heated to a temperature of 70°C, which enables the *Taq* polymerases to incorporate free dNTPs at the optimal operating temperature for *Taq* and thus synthesizing new DNA strands.

After the first cycle, the targeted genomic template has, like DNA in living systems, been replicated in a semi-conservative manner meaning that for each original double-stranded DNA (dsDNA) template instead now exists two dsDNA molecules—each one comprised of an original or "parental" template strand plus a newly-synthesized complementary or "daughter" strand. Once the first cycle ends a second cycle immediately begins repeating steps 1–3 above. A diagram showing the steps during the first two PCR cycles is illustrated in Figure 5.5. During the first cycle the template consists of long strands of genomic DNA, each of which could span thousands if not millions of bases (even entire chromosomes). However, the synthesized DNA strands during the first cycle are far shorter in length because the time for extension is too short to allow the DNA polymerase to synthesize more than about 2,000 bp of DNA. Note, a technique known as "long PCR" can copy templates that span thousands or tens of thousands of bases but that is a more specialized form of PCR, which we will discuss later in this chapter. In Figure 5.5, notice that half of the newly synthesized strands at the end of the second cycle become even shorter in length. They span the exact distance between the 5′ ends of the forward and reverse primers with the target sequence located in between the primers (Mullis et al. 1986). This is how PCR can specifically target and copy a particular genomic locus. Within a few cycles these shorter double-stranded products will be far more involved in the synthesis of products than the original genomic DNA templates. Note also that the forward and reverse primers become permanently incorporated into each newly synthesized double-stranded product. Examining Figure 5.5 again, we not only see that the original template strands are available again for DNA synthesis for the start of the second cycle, but the new strands from the first cycle *also* contain target sequences for the forward and reverse primers. This means that the number of products at the conclusion of the second cycle will double to four. Figure 5.6 shows that the number of templates doubles after each cycle showing this to be a geometric growth process (i.e., $2 \rightarrow 4 \rightarrow 8 \rightarrow 16\ldots$ or 2^n, where n = number of cycles; Saiki et al. 1988), which is why PCR is also commonly referred to as DNA *amplification*. PCR is so sensitive that with just one double-stranded template DNA to begin a PCR, one can obtain 2^{35} or 34,359,738,368 copies after 35 cycles! If there are hundreds or thousands of genomic template copies present at the start of the reaction, then there will be trillions of copies when PCR is finished!

Let's now take a closer look at the synthesis of PCR products during the extension step of a PCR after a number of cycles have already been completed. Figure 5.7a shows two complementary strands of a product generated in some previous cycle and each strand already has a primer bound to the appropriate template sites. Thus, the denaturation and annealing steps have already been executed in this cycle. Once the primers become bound to their templates and the temperature is raised to 70°C the extension step commences—the DNA polymerases can start to synthesize new strands of DNA. Figure 5.7a shows how both primers become incorporated into the newly forming product. It is also important to notice that downstream from the bound primer the newly synthesized strand contains the complementary sequence to the other primer. This means that every new strand can itself be used as a template in a future cycle of synthesis (Figure 5.7b). After extension is complete, a 540 bp long product that is double-stranded from end to end is produced (Figure 5.7b). As you can see, each product consists of two original primers plus two newly made strands downstream of each primer.

It should now be clear that PCR can readily generate billions or more copies of a specific target locus. Interestingly, the accumulation of PCR products over the course of a typical PCR (e.g., 35 cycles) does not resemble an exponential growth process, but rather is better defined by an S-shaped or logistic growth curve (Saiki et al. 1988; Innis et al. 1990) as depicted in Figure 5.8. Thus, during the first ~20–25 cycles of PCR, the rate of amplicon production is exponential. However, after this point the rate of growth starts to slow down until the growth curve flattens out into a plateau of low growth (Saiki et al. 1988; Figure 5.8). A number of explanations have been advanced to explain this "plateau effect" in PCR including depletion

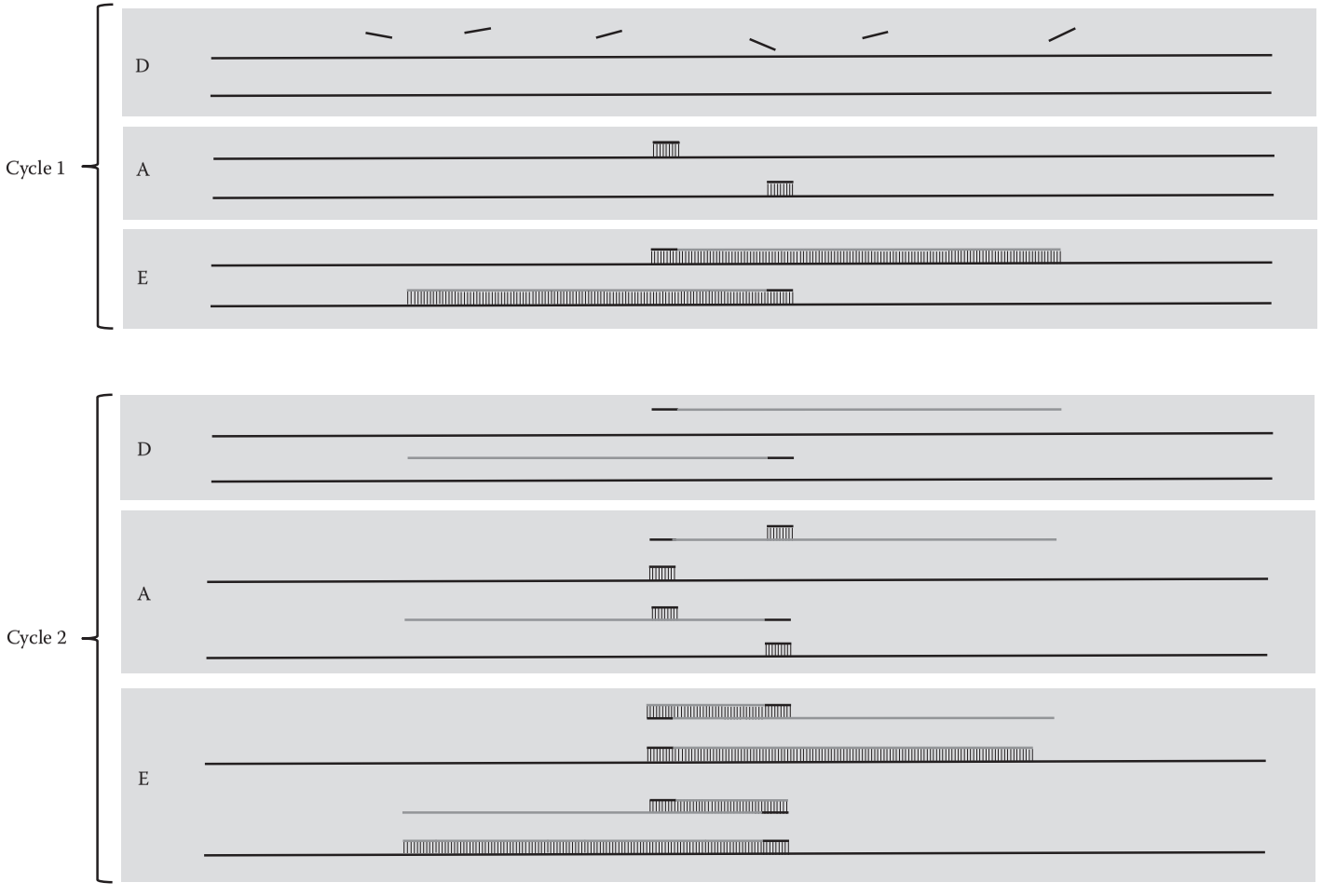

Figure 5.5. The first two cycles of PCR. Reaction starts with a genomic DNA template (black lines) and primers (short dark gray lines). D, denaturation step; A, annealing step; and E, extension step. Vertical lines depict hydrogen bonds. New strands of DNA are gray.

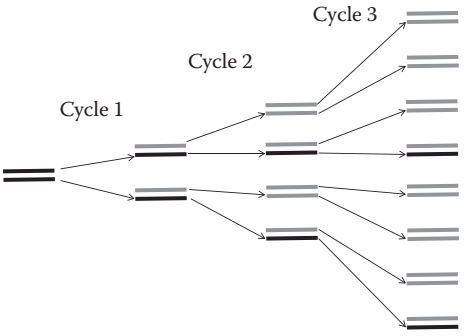

Figure 5.6. PCR replication from cycles 1 through 3 illustrating the geometric increase in dsDNA. Cycle 1 begins with one double-stranded template shown as two black (parental) strands and ends with eight after cycle 3. New (daughter) strands are gray.

of reagents such as dNTPs, primers, and Mg^{2+} cations; inactivation of *Taq* polymerases due to the high heating for many cycles; and inhibition of *Taq* caused by the accumulation of end products such as pyrophosphates and target products (Innis et al. 1990; McPherson and Moller 2006). Although all of the aforementioned possibilities could have some impact on amplicon production in PCR, the study by Kainz (2000) showed that this slowdown can simply arise due to negative feedback (i.e., density dependence) provided by the accumulation of the amplicons themselves. Regardless of its cause, however, the plateau effect in PCR seems to not be a concern. This does not mean that investigators should never consider fine-tuning their PCRs by fiddling with reagent concentrations or mixing in additives. Such fine-tuning, which is generally referred to as "PCR optimization," especially when new primers are being tested is actually a key part of the PCR experience. We will consider PCR optimization throughout the remainder of this chapter and in Chapter 8 when we consider primer design and loci development.

5.3 PCR PROCEDURES

Like agarose gel electrophoresis, the basic equipment and methodology of PCR has changed little

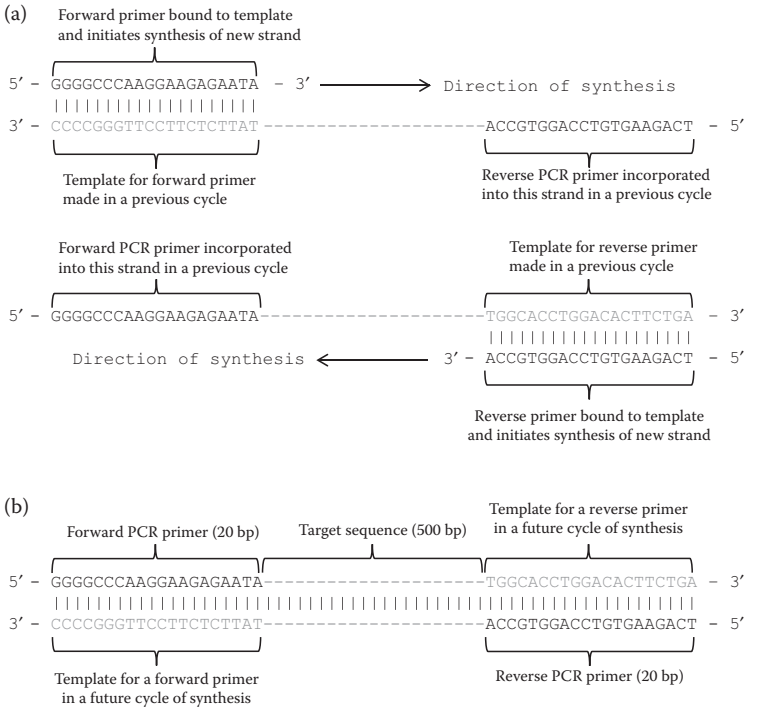

Figure 5.7. Synthesis and anatomy of a 540 bp PCR product. (a) Synthesis of two products from two complementary strands during the extension phase (after many cycles). (b) Parts of a PCR product. For clarity the target sequence is indicated as dashes. Note, most of target sequence not shown due to lack of space. Vertical bars indicate hydrogen bonding between complementary bases. Gray represents strands of DNA synthesized during PCR and primer bases are shown as black letters.

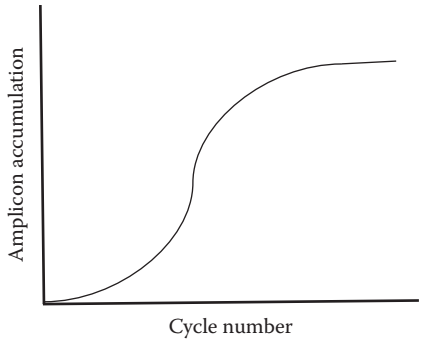

Figure 5.8. Growth curve depicting the accumulation of PCR amplicons during a PCR.

over the past 30 years. From start to finish, a PCR experiment requires 3–6 hours of lab time. There are three steps to conducting a PCR experiment: (1) *setting up the reactions*; (2) *thermocycling*; and (3) *agarose gel electrophoresis*. If a PCR is found to have failed to some degree, then PCR troubleshooting will be needed.

5.3.1 Preparation of PCR Reagents and Reaction Setup

The first step to conducting a PCR experiment is to prepare the PCR reagents prior to setting up the reactions. Before we discuss the processes of combining all reagents into the reaction tubes, thermocycling, and electrophoresis we will now review basic details about preparing each PCR reagent.

5.3.1.1 PCR Reagents

Primers—We are now familiar with the function of forward and reverse PCR primers in the PCR, but where do these primers come from and how are they designed? This is a complicated subject thus for now we will assume that the appropriate primers are already in hand (previously published or designed by someone else). We will explore the topic of primer design in Chapter 8.

Commercial manufacturers of primers such as Integrated DNA Technologies (IDT) desalt and dry (i.e., lyophilized) newly synthesized primers prior to shipment to customers. When the new primers arrive in the lab, they must be prepared for use in a PCR. The first step typically involves resuspending the dried primer with molecular biology grade water or TE buffer to make a 100 μM "concentrated" stock solution. The next step is to use the concentrated stock to make a more dilute 10 μM "working" stock solution for direct use in PCR. Although 10 μM working stocks will work for most routine PCRs, some researchers prefer to use less concentrated working stocks; so different working stocks can be made accordingly using the concentrated stock solution. The main reason why a concentrated stock is necessary is that repeated freeze–thawing of working stocks will lead to degradation of the primers, which in turn will diminish the efficiency of PCR. Thus, fresh working stocks of primers should always be used in PCR. To avoid contamination of the primer stocks it is highly advisable that this work be performed in a relatively sterile space such as a laminar flow hood and using aerosol-prevention (usually called "barrier" or "filter" tips) pipette tips.

Although making concentrated and working stocks of primers are simple procedures, it is critically important that this work be correctly done. The following is an example protocol for mixing up concentrated and working stocks of primers manufactured by IDT that my students and I have successfully used:

Step 1: Calculate the volume of molecular biology grade water or TE that you will add to each tube containing dried primers. For each primer you will need to look up the amount of primer synthesized. On the datasheet that comes with each primer, under the heading "amount of primer," you will see the quantity of primer in units of nmoles that was synthesized. This number is often between 20 and 40 nmoles. Multiply this number by 10. This new number will be the volume of water or TE in microliter that you will add to the primer tube. For example, if the amount of primer is 25 nmoles, then you will need to add 250 μL of sterile pure water or TE to the tube of dried primer. However, before opening the new primer tubes to add the water or buffer, briefly centrifuge the tubes for 1 minute in the microcentrifuge at maximum speed in order to guard against losing dislodged primer pellets.

Step 2: Add the correct volume of water or TE to each primer tube using a pipet with a filter tip.

Step 3: Allow the tubes to sit for 10 minutes at room temperature so the DNA pellet goes into the solution.

Step 4: Use the pipet to gently mix the primer solution (by pipetting up and down slowly about 20 times) to make homogeneous. If the correct amount of water or TE was added, you should now have a 100 μM solution of the primer. This concentrated primer stock is now ready for use to make diluted primer for PCR (i.e., the primer working stock).

Step 5: Now label a set of clean and sterile 1.5 mL microcentrifuge tubes (one for each primer) by writing the name of the primer and concentration (10 μM) on the lid. Write the date on the side of the tube. These new tubes will be the diluted primer ready for PCR.

Step 6: Add 90 μL of molecular biology grade water to each of the newly labeled microcentrifuge tubes. Next, add 10 μL of concentrated primer to each microcentrifuge tube. Use the pipet to gently mix the solution to make a homogenous concentration of the primer. You now have 100 μL of 10 μM primer ready for PCR. Store the primer working stock tubes at −20°C when not in use.

dNTPs—Deoxynucleoside triphosphates or "dNTPs" (i.e., dATP, dTTP, dGTP, and dCTP) are required for a PCR. As with primers, fresh dNTPs perform better than well-used solution of dNTPs due to the negative effects of repeated freezing and thawing of the nucleotides. A 10 mM solution containing equal molar amounts of each nucleoside triphosphate is often used in a PCR setup. Ready-to-use mixes containing all four dNTPs can be purchased from biotech supply companies or you can obtain the four types of dNTPs in separate tubes and then easily make your own mix.

Taq polymerase—In addition to adding the *Taq* enzymes to the reaction mixture, a 10× magnesium-containing buffer (i.e., MgCl) must also be added because, as we saw earlier, *Taq* polymerase requires free Mg^{2+} cations in order to catalyze the nucleotide polymerization process.

Molecular biology grade water—Critical to the success of PCR is the use of sterile, pure, and deionized water in all lab procedures otherwise problems can arise later. For example, DNA can be degraded or destroyed by DNA-destroying enzymes or chemicals may inhibit PCR. Such "PCR" water can be made in a lab by proper distillation and autoclaving procedures or can be purchased from life science or chemical supply companies as "Molecular Biology Grade Water," "HPLC Water," etc.

Template DNA—For most PCRs, DNA extracted from tissues is later used as template for the PCRs. Often this extracted DNA will be too concentrated for use in PCR so a sample should be diluted in the reaction. There are several advantages to diluting a sample of extracted DNA for use in PCR. These include diluting any PCR inhibitors (e.g., DNA polymerases are sensitive to salt) present in the original extraction; avoiding overamplification of the target locus; avoiding amplification of nontarget regions ("PCR artifacts"); and conserving DNA template by not using more than is required for a successful amplification. The old adage "less is more" applies well to the problem of how much template to add to a PCR. Although using a sample of template that is too dilute might lead to poor amplification of the target gene, the advantages to diluting the original template outweigh the disadvantages. So, how much template DNA is needed for each reaction? Since most PCR protocols require only 1.0 μL, the important thing to consider is the concentration of the extracted DNA. Although PCR can be robust to template concentration, 10–100 ng/μL represents an ideal range. It is very advantageous to start with a concentrated solution of extracted DNA and then dilute it to 10–15 ng/μL for use in PCR, as more than enough template will be available for a strong amplification while any inhibitors present in the original solution are less effective because of the dilution.

5.3.1.2 Importance of Making Reagent Aliquots

Researchers should always make aliquots of all working stocks of PCR reagents. This includes primers, molecular grade water, master mix (see Section 5.3.1.3), and template DNA. Usually such aliquots are dispensed into 1.5 mL microcentrifuge tubes then are stored in a −20°C freezer. One reason this is a good idea is that some of these reagents (especially the primers, dNTPs, and template DNA) become more degraded each time they are frozen and thawed making them perform more poorly in PCR. A second reason is that making aliquots greatly minimizes the risk of a "global" PCR contamination event in a laboratory. By using aliquots rather than

concentrated stocks, contamination outbreaks can be confined to particular aliquots or PCR experiments. It only takes one misplaced pipette tip to contaminate an entire concentrated stock rendering that reagent useless for future PCRs, not to mention the wasted lab resource and time lost doing contaminated PCRs! Thus when new reagents arrive to a laboratory, a good strategy is to immediately go to a clean space such as a laminar flow hood and use filter tips to make aliquots. Afterward, the aliquots should be transferred to a freezer where they are stored. Be sure to also write the date the aliquot was made on the side of the tubes so that each lot can be monitored for quality control.

5.3.1.3 Setting Up PCRs

Typically PCRs are performed in 0.2 mL plastic PCR tubes or 96-well plastic PCR plates (0.2 mL wells). While performing this step, it is good to keep all reagents and tubes (or plates) on ice, which will help preserve the reagents. Once the reagents are thawed, they can be pipetted into the reaction tubes.

Negative and positive PCR controls—In addition to setting up a PCR for each of your different template DNA samples you should also set up a *negative control* and a *positive control*. The negative control is a PCR setup that contains all reagents except any template DNA—water is substituted for DNA. This control is important to every PCR experiment because it will indicate a DNA contamination problem when it exists. If all reactions containing template DNA plus the negative control amplify a product, then this indicates that foreign DNA (usually PCR products) contaminated one or more of the PCR reagents. The positive control contains all the reagents plus a template DNA sample that has successfully worked in prior PCR experiments with the same primers that are being used. Thus the positive control uses a proven DNA template sample. If the positive control amplifies correctly but the other samples do not, then you can safely conclude that the other DNA templates are somehow defective. If none of the samples containing DNA including the positive control amplify, then there are multiple possible reasons why the PCR failed. We will discuss this again when we consider troubleshooting failed PCRs later in this chapter.

Making a PCR master mix—When you are ready to set up the PCRs the first thing you will do is to retrieve aliquots of each reagent from the freezer and then thaw them. It is best to thaw the reagents and perform the pipetting steps inside a laminar flow hood because working in a relatively sterile environment will help to reduce the chances of contamination. Once the reagents are thawed, the next step is to make a "master mix," which is a mixture prepared in a 1.5 mL microcentrifuge tube containing exactly the right amounts of primers, dNTPs, enzyme buffer, water, and *Taq* (i.e., all components except the DNA template). Such master mixes are useful because they save you lots of time and pipette tips and reduce the likelihood of pipetting errors. Although micropipettes tend to perform well, you should expect that one or more of your pipettes, especially the larger volume pipettes (e.g., 200 µL and 1,000 µL models), may not deliver the exact desired volume to your master mix or PCR tubes. This potential problem is easily overcome by adding in extra reagents when making your master mix. For example, by making 10% more than you theoretically need to accommodate all your samples, you ensure that you will not run short on master mix. An example recipe for making a master mix for a PCR experiment that includes 30 different template DNA samples plus negative and positive controls is shown in Table 5.2.

Note, you can save even more time (fewer pipetting steps) and consumables (use fewer pipette tips) by using a so-called "pre-made master mix" not to confuse with the "master mix" discussed earlier. A *pre-made* master mix is usually a large stock solution that contains dNTPs, enzyme buffer, *Taq*, and water to set up hundreds or thousands of PCRs; thus it is a large stock for a great many PCR setups. Pre-made master mix usually does not include primers, though if your laboratory primarily uses a single primer pair (e.g., for DNA barcoding), then making a pre-made master mix that includes primers may be worthwhile.

There are two options for obtaining pre-made master mixes: purchase from a biotech supply company (e.g., Promega) or make your own. Commercially available pre-made master mixes are convenient because they save you time and trouble, are easy to use, and usually produce good PCR results. Also, some companies offer pre-made master mixes that already contain a loading dye,

TABLE 5.2
Example showing the setup for a PCR experiment that includes a total of 32 PCRs, which includes 30 DNA samples plus positive and negative controls

Master mix (MM) reagents	Volume/reaction (μL)	# Reactions[a]	MM volume (μL)
Water (molecular biology grade)	10.5	35.2	369.6
Taq buffer (10×)	1.5	35.2	52.8
dNTP mix (10 mM)	0.9	35.2	31.7
Forward primer (10 μM)	0.5	35.2	17.6
Reverse primer (10 μM)	0.5	35.2	17.6
Taq polymerase (5 U/μL)	0.1	35.2	3.5
Totals	14.0	35.2	492.8

[a] Signifies that the # of reactions to use in the calculations = actual # PCRs to be set up (i.e., 32 reactions) + extra reagent (0.1 × total number of reactions) to account for minor pipette errors. Thus enough reagents to accommodate an additional 3.2 PCRs should also be included to ensure that enough master mix is made.

which simplifies the loading of PCRs into agarose gels. A good alternative to the commercial pre-made master mixes, is to purchase the individual reagents separately and make your own pre-made master mix. The main advantage is you can lower your reagent costs by buying the components (dNTPs, *Taq*) separately. Another great advantage is that you have more flexibility on which type of DNA polymerase you can use. Biotech supply companies offer a variety of DNA polymerases suitable for PCR to choose from, ranging from the inexpensive (standard) to the expensive high-performance polymerases. My students and I have found that making our own pre-made master mix using a higher quality *Taq* polymerase produces better PCR results at a lower cost than buying the commercially available pre-made master mix. Testing a variety of different formulations of pre-made master mixes involving different *Taq* polymerases could provide a big payoff to the individual or laboratory by reducing reagent costs while improving PCR performance at the same time.

Returning to the act of setting up the reactions, once the reagents are thawed and ready to be used be sure to invert each reagent tube several times (or gently pipet) to mix the contents—this is especially critical for the pre-made master mix. You are now ready to begin pipetting the reagents into the PCR tubes. Use extreme care when pipetting to avoid cross contaminating samples, reagents, etc. When you are finished pipetting and have sealed the PCR tubes with their caps, it is time for the thermocycling step.

5.3.2 Thermocycling

Once the reactions are set up they are then ready to be placed in the thermocycler to start the PCR. Thermocyclers have computer touch pads so the operator can specify the actual PCR profile to be used. An example profile would be:

Step 1: 94°C for 30 seconds (initial denaturation step)

Step 2: 94°C for 30 seconds (denaturation step)

Step 3: 50°C for 30 seconds (primer annealing step)

Step 4: 70°C for 60 seconds (extension step)

Repeat Steps 2–4 total of 35 times

Although innumerable varieties of PCR profiles have been successfully used, the basic one shown above can work well for many applications assuming the correct annealing temperature is specified. This is because the temperature for denaturing the DNA (Steps 1–2) and extension of newly created DNA strands (Step 4) are standard for the vast majority of PCR experiments. On the other hand, the annealing temperature (Step 3) can vary between approximately 45–65°C depending on the optimal annealing temperatures for the primers being used. Here, we are assuming that 50°C is the optimal temperature for the primers. Other important variables in this profile include the time spent at Steps 2–4 as well as the total number of cycles to be run. For many PCRs a 30-second denaturation time is adequate

to convert the double-stranded template into single-stranded DNA. However, some researchers prefer to include one long (2–5 minutes) denaturation only during the first step to ensure that all genomic DNA has been converted to single-stranded form thereby generating the maximum amount of single-stranded templates for the primers. Regarding the extension step, 30 seconds will be adequate time for the polymerase to complete a new strand of DNA provided the product size is less than about 500 bases, but one minute should be used for products that are 1,000 to 1,500 bp (Palumbi 1996). Typically, 35 cycles are adequate. However, if in doubt about which program to use, then you should consult the guidelines established by the manufacturer(s) of the PCR reagents you are using or see Palumbi (1996) for optimization strategies.

5.3.3 Checking PCR Results Using Agarose Gel Electrophoresis

In Chapter 4 on DNA extraction, we learned about using agarose gel electrophoresis to check the results of a DNA extraction. This same method also plays a key role in PCR because it is used to determine whether or not the PCRs functioned as expected. If the PCR worked well, then the investigator will only see the ladder and a single bright band in each well of the gel corresponding to the expected product sizes and with no evidence of contamination (Figure 5.9). If success is achieved, then the PCR products will be ready to sequence. However, if some degree of failure is experienced, then the investigator must perform some troubleshooting in order to hopefully fix the problem(s).

The same reagents and methods used to make agarose gels discussed in Chapter 4 remain the same for evaluating PCR products. The main difference between extraction and PCR gels, is that expected PCR product sizes are usually less than 1 kb, therefore a more concentrated gel (2%) is often used. This is because the more concentrated gel will provide greater resolving power for smaller linear fragments of DNA than a less concentrated one. Also, as before, the ladder functions as a positive control for the electrophoresis steps as well as enable the sizing of products. Typically, 5 μL of PCR product plus 1 μL of 6× loading dye are loaded into each gel well.

5.4 PCR TROUBLESHOOTING

When examining the results of a PCR experiment on an agarose gel the first thing you will be eager to look for will be the single bright bands that correspond to your expected product sizes, such as in the example shown in Figure 5.9. The presence of bands corresponding to their expected sizes usually means that specific amplification of the target locus has been achieved. If your sample lanes show good products and the negative control lane shows no evidence of DNA contamination (i.e., no band in this lane), then you are ready to prepare the products for sequencing. However, if the gel reveals any problematic issues with your PCRs such as missing target bands, weak target bands, multiple bands, smears, etc., then you will need to perform a troubleshooting analysis to determine the causes of the failures.

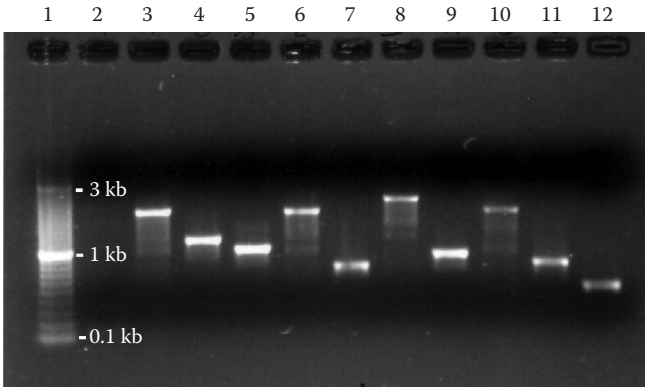

Figure 5.9. Photo of 2% agarose gel stained in ethidium bromide. Lane 1 is ladder; lane 2 is a negative control; and lanes 3–12 are PCR products for ten different target loci. Dark oval areas (below each number) are sample-loading wells in the gel.

Given that the entire process beginning with DNA extraction and ending with agarose gel electrophoresis of your target PCR products involves a number of procedures and intricate chemical reactions, numerous opportunities exist for one or more problems to arise causing the PCR experiment to fail in one way or another. Although the basic PCR procedure is straightforward, sooner or later everyone obtains gel results that indicate something went wrong somewhere during the PCR or electrophoresis procedures.

PCR gels can have many different appearances depending on the nature of the PCR or electrophoresis problems and they can therefore reveal important clues for troubleshooting problems. The first thing to rule out is a problem with the gel itself because if something malfunctioned with the gel electrophoresis procedure, then you will not be able to evaluate your PCRs. Nucleic acid stains such as ethidium bromide tend to lose their activity over time, especially if they are exposed to light. Thus, whenever a PCR gel does not show any DNA bands—not even the ladder, then the possible culprits include defective electrophoresis reagents or procedures. Other gels may show ladder and PCR bands but the bands are feint and difficult to see. This latter gel result suggests that the PCR worked but that the weak stain makes the bands difficult to visualize. A new gel with fresh nucleic acid stain should be prepared and run to retest the PCR products. If you can rule out gel problems and the PCR results are (still) disappointing, then you will need to troubleshoot the PCR itself. A summary of the most common PCR problems is shown in Table 5.3. Each of these scenarios is discussed in more detail including troubleshooting tips.

Consider another possible PCR gel result in which the only DNA bands on the gel are within the ladder lane—no bands in the positive control or sample lanes. In this scenario, there is sufficient evidence to conclude that the PCRs must have failed. Although the template DNA could be to blame (of course if any had worked successfully before, then this can be ruled out), because the positive control also did not work indicates that a more global problem exists (e.g., defective master mix or thermocycler). Although the thermocycler could be the problem, this is rarely the case as it is far more likely that the problem exists with one or more components in the PCR. If using a commercially made master mix, then you could set up a new PCR experiment but this time substitute in a new master mix. If the master mix was made in the lab then you need to troubleshoot the exact cause (bad *Taq*?, bad dNTPs?, etc.). This can be done by setting up a new PCR experiment in which all tubes get the *same* template DNA (use the proven positive control DNA), but each one differs by a single experimental variable:

Tube 1: *New Taq*, old dNTPs, old water, old forward primer, old reverse primer

Tube 2: Old *Taq*, *new dNTPs*, old water, old forward primer, old reverse primer

TABLE 5.3
Troubleshooting PCR experiments using information obtained from agarose gels

Ladder visible	Band(s) in (−) control	Band(s) in (+) control	Band(s) in sample lanes	Likely cause(s)
No	No	No	No	Defective ladder, stain, or gel[a]
Yes (D)	No	Single, correct size (D)	Single, correct size (D)	Weak stain
Yes	None	None	None	Defective master mix or thermocycler
Yes	Yes	Yes	Yes	Contaminated master mix
Yes	None	Single, correct size	None	Problem with template DNA
Yes	None	Single, incorrect size	Single, incorrect size	Nonspecific amplification
Yes	None	Multiple bands	Multiple bands	Nonspecific amplification
Yes	None	Smear	Smear	Over-amplification of target

NOTE: Each row summarizes observations in the form of presence or absence of amplified DNA ("bands") from a particular gel and identifies the likely cause(s) for those results. "D" means band was difficult to see on the gel.

[a] Indicates that problem could be electrophoresis conditions.

Tube 3: Old *Taq*, old dNTPs, *new water,* old forward primer, old reverse primer

Tube 4: Old *Taq*, old dNTPs, old water, *new forward primer,* old reverse primer

Tube 5: Old *Taq*, old dNTPs, old water, old forward primer, *new reverse primer*

In this case, a "new" reagent is either a new unopened product or one that is freshly remade using a concentrated stock solution, whereas the "old" reagent is the same reagent stock used in the prior failed PCR experiment. This is an easy way to troubleshoot many PCR problems because it is usually only one defective reagent that causes PCR failures. If none of these tests reveals the problem, then you can make a sixth tube that is made using all new reagents. Although it is quicker and less trouble to just throw away all old PCR reagents and start again with fresh ones, taking the time to troubleshoot these problems may help you avoid the same problems in the future thus saving you more time in the long run.

On other gels you might see that all of the samples including the positive control (if one is used) as well as the negative control amplified, a result that implicates DNA contamination. Pipetting errors or using pipettes contaminated with DNA (especially PCR products) is the usual culprit. If a contamination problem exists, then sterilizing the pipette(s) using UV light and following the approach outlined above for troubleshooting failed PCRs can help you understand the nature of the contamination problem. Taking the time to identify the source of contamination is a key step forward toward preventing similar problems in the future.

Although a positive PCR control is optional—provided the PCRs have been working as expected in previous experiments, this control can nonetheless be useful if the sample tubes show little or no sign of amplification. Remember, if the positive control amplifies, then this immediately allows you to rule out many possible causes for the failed PCRs—master mix, thermocycler, and gel electrophoresis. If the positive PCR control sample is successfully amplified, then this tells you that everything worked as expected except for the DNA templates in the samples. This would then point to the DNA obtained from the extractions. Perhaps those samples were too degraded to be used in PCR or something else was wrong with the template DNA (e.g., inhibitors or low template concentration). Occasionally, DNA templates contain too much salt or ethanol leftover from the extraction process, which can lead to inefficient or no amplification because these are inhibitors of the *Taq* enzyme. Recheck the template concentration and quality (extraction gel and UV spectrophotometry results). If a DNA extraction gel shows that the DNA is both concentrated enough for PCR and mostly undegraded, then diluting the template further and/or desalting the template using ethanol precipitation (with at least two 70% ethanol wash steps and ensuring that all ethanol is removed prior to the last step) can lead to dramatically better PCR results.

Other problems can include smearing of DNA or multiple bands within gel lanes. If the positive control amplified correctly but the samples did not, then this is likely a template problem; specifically the template could have been too concentrated leading to over-amplification of the target band. If the positive control and sample lanes all show smears or multiple bands, then this suggests that nonspecific amplification occurred during the PCR leading to multiple genome targets being amplified. One possible remedy for this is to increase the annealing temperature in 1–2°C increments until the optimal annealing temperature is identified. This can be accomplished in a single PCR experiment using a "gradient thermocycler," in which case a range of annealing temperatures can be simultaneously used or, if using a nongradient thermocycler, you can perform a number of separate experiments, each one using a different annealing temperature. Sometimes this problem cannot be resolved by trying different annealing temperatures. If multiple bands including a single discrete target band are evident in a gel, then you can use a gel purification protocol to obtain the desired PCR products. Alternatively, another method is to rerun the PCR products on a "low-melt" agarose gel. Then use a sterile pipette tip to obtain an agarose plug taken from the center of the target band on the gel. Next, place the agarose plug into a 1.5 mL microcentrifuge tube containing 1 mL of molecular biology grade water then heat the tube to a high enough temperature (i.e., 65.5°C) to melt the agarose plug. Use 1 μL of this solution as your DNA template (instead of the original DNA extraction) in a new PCR. This can be a surprisingly simple and fast method for acquiring PCR products though safe procedures (especially for eye protection) for working with

a UV illuminator should always be carefully followed. Note that these gel-based methods for obtaining target PCR products are labor intensive and slow and thus they are not practical methods for obtaining many PCR products. If large numbers of products are needed, a better strategy is to find alternative primers or redesign them to function better.

5.5 REDUCING PCR CONTAMINATION RISK

A negative PCR control only tells you whether a contamination problem exists. Good preventative measures in the laboratory are needed in order to minimize DNA contamination risks. Contamination in the DNA laboratory increases the amount of time and costs of consumables not to mention the frustration on the part of the researcher. Establishing preventative measures in the lab can dramatically diminish the incidence of contamination.

DNA contamination often occurs when concentrated DNA such as freshly extracted genomic DNA and *especially* PCR products contaminate some component of the PCR (reagents, pipette, etc.). Although a pipette can become contaminated by direct contact with concentrated DNA while the pipette is being used to transfer DNA from one tube to another, concentrated DNA can also become *aerosolized* thus making cross-contamination of samples in adjacent tubes even more likely. DNA contamination can afflict a variety of things in the lab including DNA extractions, plastic reaction tubes, pipette tips, pipettes, reagents for PCR, etc. Although the source of a DNA contamination problem can be from genomic DNA, PCR products and plasmids represent the usual sources of contamination. This is because PCR products and plasmids tend to be highly concentrated, whereas genomic DNA is far less concentrated. PCR products and plasmids are often referred to as *high copy* DNA while genomic DNA is called *low copy* DNA. Given that PCR can amplify DNA even from only a few copies of template DNA, it should be easy to see that this process is quite vulnerable to DNA contamination and therefore implementing measures to prevent contamination of a PCR *before* conducting PCR is of critical importance.

By recognizing the threat posed by high copy DNA, a number of things can be done to prevent PCR contamination. Earlier we already discussed the importance of making aliquots of PCR reagents, which can greatly reduce the incidences of contamination. However, there are additional measures that can be taken. Many DNA laboratories maintain separate "Pre-PCR" and "Post-PCR" spaces so that DNA extractions, PCR setup, and other pre-PCR procedures are performed away from areas that are likely already contaminated with high-copy DNA such as thermocyclers and gel electrophoresis stations (see Figure 1.3).

Pre-PCR area—This area of the lab should have its own dedicated set of pipettes and consumables. Filter tips can further reduce the contamination risk because the filter contained within the plastic tip acts as an effective barrier to DNA thus helping to avoid cross contamination. A laminar flow hood can also help reduce contamination problems by providing a well-lit clean space where DNA extractions and PCR setup can be performed. As we discussed in Chapter 4 a 10% bleach solution will destroy DNA so occasionally wiping down all counter-tops can help maintain a hygienic work environment. Many plastic consumables such as microcentrifuge tubes and PCR plates have been pre-sterilized in the factory so you may not need to autoclave them. However, if you have any doubts about the sterility of your plastic consumables, then you should autoclave them or use a UV cross-linker to be sure (except filter tips). Although DNA contamination can arise despite using all precautions (e.g., pipetting error), the incidence of contamination can be greatly contained. This will help prevent a more global contamination problem that would lead to the loss of valuable DNA extractions and other reagents. In addition to protection from contamination, making aliquots can reduce the freeze–thaw damage to the original or concentrated stocks caused by repeated (and needless) retrieval of solutions from freezers. Concentrated stocks of extracted DNA, primers, etc. should be stored and used with utmost care.

Post-PCR area—DNA contamination becomes less of a worry once PCR has already been performed because the remaining steps toward obtaining sequence data are not likely to be compromised by template DNA contamination. Although Sanger sequencing is a PCR-like procedure (Chapter 6), the cycle sequencing reaction is robust to minor PCR-product contamination because the presence of the correct PCR template will outnumber the foreign template and therefore good quality sequences can still be obtained assuming the

primers work equally well on all templates. The main requirements of a post-PCR area include a dedicated lab space where the thermocycler(s) are located as well as a dedicated set of pipettes, and both filter and nonfilter tips. The filter tips should be used when performing agarose gel electrophoresis to check the results of genomic DNA extractions, whereas nonfilter tips (which are less expensive than filter tips) can be used for loading samples of PCR products on gels.

5.6 HIGH-THROUGHPUT PCR

High-throughput PCR often involves the use of plastic "microplates" with 96 separate 0.2 mL wells arranged in 8 rows by 12 columns; the positions of samples in the plate can be noted using the row letter (A to H) × column number (e.g., C7), which is important when dealing with 96 samples on one plate. Moreover, it is important to distinguish between two varieties of microplates: "un-skirted" and "skirted." Laboratories may need both types of microplates because some thermocyclers only accept the former, whereas it is safer practice to use the latter type of microplate to send PCR products by mail (though you should read the recommendations given by the sequencing laboratory that will be receiving your plate). As an alternative to using plates, some flexibility in the numbers of samples processed can be added simply by using individual 0.2 mL PCR tubes or especially 0.2 mL "strip-8" tubes. If single tubes are arranged in "8s" on a plastic 96 tube rack, or if strip-8 tubes are similarly arranged on a rack, then an 8-channel pipette can be used to great advantage. Such an approach can be considered "quasi high-throughput."

5.6.1 Setting Up PCRs in a 96-Sample Microplate Format

The master mix should be poured into a clean (sterile) multichannel plastic tray or distributed among eight 0.2 mL PCR tubes situated in a rack. Doing this will allow you to use the 8-channel pipette to quickly transfer the master mix to each well in the new PCR microplate. Next, the same pipette is used to transfer the diluted DNA template from the template plate (or tubes) to the PCR plate. When the master mix and template have been added to the PCR plate, great care should be used to seal the plate so that liquid cannot escape. Some labs prefer to use silicone rubber "mats," but be careful that the lid of your thermocycler can apply sufficient pressure to seal each well in the plate otherwise some of the reaction mixtures may escape from the wells during the thermocycling process. An excellent alternative method to seal the plate is to use strip-8 tube caps (i.e., each 8-sample column is sealed with a different strip of plastic caps). Following thermocycling be sure to perform the following two procedures with your PCR plate before unsealing the wells. First, cool the plate in a refrigerator or freezer. Hot samples have a greater chance of volatilizing PCR products, which not only results in losing a portion of your samples, but also can cause cross-contamination problems. Secondly, spin the plate in the tabletop centrifuge for 1 minute in order to push all liquid to the bottom of each well in the plate, which will minimize the chance of liquid escaping from the plate upon opening.

By now you have realized that anytime you can switch from a single-channel micropipette to using a multichannel micropipette to handle a large number of samples you are going to greatly economize your lab work. As we will see in Chapter 6, there are many advantages to conducting your PCRs in the 96-sample format, as not only does it reduce the time and per sample costs for PCR, but it facilitates downstream applications such as DNA sequencing.

5.7 OTHER PCR METHODS

We conclude this chapter by considering several other types of PCR, which have important applications in phylogenomics. The first variant of PCR we will discuss is called *hot start PCR*. This procedure can improve the specificity of problematic PCRs by preventing the formation and coamplification of short nontarget PCR products as a result of primer–primer interactions, which can lead to poorer DNA sequencing results. Another variant of PCR, called *long PCR*, is a technique for amplifying loci up to 40 kb long. Although long PCR has thus far had limited applications, it has proved useful for amplifying and sequencing whole mitochondrial genomes (Zardoya and Suárez 2008). This technique may also be useful for avoiding mitochondrial pseudogenes or "numts" (Sorenson and Quinn 1998). The third variety of PCR we will discuss is called *RT-PCR*, which has been an important method for phylogenomic

studies concerned with the evolution of protein-coding genes.

5.7.1 Hot Start PCR

The best-performing PCR primers only amplify the target product, a result that usually guarantees the acquisition of high-quality DNA sequence data. However, occasionally researchers use primer pairs that amplify the target product plus unwanted *primer dimers*.

Primer dimers are short (~40–60 bp) double-stranded PCR products that arise because of design defects in the primers that cause them to hybridize to each other during PCR. Because duplexed primers can serve as both template and primer, this can result in the accumulation of a very large number of these products. Primer dimers interfere with the sequencing of the target products and therefore diminish the quality of sequence data. We will take a closer look at why primer dimers form and how to prevent their occurrence in routine PCR when we discuss primer development in Chapter 8.

Although it is possible to obtain useable DNA sequences from PCR products containing primer dimers, it is preferable to use some method to remove and discard them from already-completed PCR products or to prevent their formation during the PCR. We will discuss PCR product purification methods in Chapter 6 but the best strategy for dealing with the primer dimer issue is to prevent their formation in the first place. This can be accomplished through careful primer design or by using a method known as "hot start PCR" (Chou et al. 1992).

Chou et al. (1992) realized that primer dimers initially form during the first cycle of a PCR—as the thermocycler begins increasing the heat toward the initial denaturation step, the problematic primers begin to anneal to each other, which can allow the polymerase to start synthesis using primers as templates. This creates a problem because during the first annealing step in cycle 1 not only will the primers anneal to the correct target sequences, but also to a far larger number of primer-based templates. The result of such a reaction will be to coamplify both target and primer dimers. In other words, the primer dimers form independently of the genomic DNA templates also present in the mixture. To address this problem, Chou et al. (1992) proposed a simple method for preventing the formation of primer dimers. The method basically involves combining all PCR reagents only after the thermocycler has reached a temperature in the 60–80°C range prior to the first denaturation step. Chou et al. used wax as a barrier to separate the enzyme (*Taq* polymerase) from the other reagents in the PCR tubes. Thus, when the high temperature is reached, the wax melts allowing for mixing of all reactants at a temperature above which short DNA molecules can anneal to other DNA molecules. The wax-based method allows for the hot start method to be more automated. More recently, biotech manufacturers have developed another such automated hot start format that uses a specially designed "hot start DNA polymerase." The hot start polymerase has an antibody attached that prevents it from synthesizing DNA at lower (i.e., below-annealing) temperatures; when the thermocycler reaches the initial denaturation temperature, the antibody is inactivated and the polymerase becomes activated. It is important to carefully read the manufacturer's instructions for hot start polymerase because longer-than-normal exposure times at 94–95°C are required to activate the polymerases. If the researcher wishes to only use the hot start method for one to several samples, then a "manual hot start" method can be implemented. This approach involves combining all reagents in the PCR tubes except for the *Taq* polymerase. The tubes are then placed into the thermocycler and heated up to the initial denaturation temperature. When the machine reaches a temperature between 60 and 80°C the researcher pauses the thermocycler, opens the tubes, and pipettes the *Taq* polymerase into each tube one at a time. The thermocycler program is then resumed and allowed to complete its normal run. Note that there are safety risks involved when attempting to access the PCR tubes while they are inside the running thermocycler because the metal plate containing the PCR tubes (or microplate) will be hot enough to cause severe burns if touched by the researcher.

The hot start method is effective at preventing the formation of primer dimers. However, if after using hot start short double-stranded "primer dimers" are still observed in the agarose gel, then the "primer dimers" are likely not primer dimers. Instead, the nontarget band(s) are likely short nonspecific targets amplified from the genomic DNA template. If the small nontarget band is

also observed in the negative control, then this is also evidence for a primer dimer problem (i.e., because primer dimers form independently of genomic DNA templates).

5.7.2 Long PCR

The vast majority of PCRs involve amplifications of loci that are <2 kb in length. One reason for this is simply due to the limitations of Sanger sequencing. As we will see in Chapter 6, it becomes more complicated and difficult to sequence PCR products that are longer than 1–1.5 kb. Another reason why most PCRs involve shorter loci is because longer loci tend to be more difficult to amplify than shorter loci. This difficulty can be encountered when using degraded genomic templates. However, even if you are using undegraded templates, the maximum length locus that can be reliably amplified using *Taq* polymerase is around 5 kb (Barnes 1994).

Unlike a number of other DNA polymerases, *Taq* polymerase lacks detectable $3' \rightarrow 5'$ exonuclease or "proofreading" activity (Tindall and Kunkel 1988; Korolev et al. 1995). Consequently, when *Taq* commits a misincorporation (base substitution) error, which on average occurs once every 4,000–5,000 bases synthesized (Innis et al. 1988; Saiki et al. 1988; Keohavong and Thilly 1989), these enzymes are unable to repair the mistakes themselves. Accordingly, one problem with using *Taq* is that PCR products generated with this enzyme can have base substitutions. However, regarding attempts to amplify longer fragments, the primary significance of these errors is that after such misincorporation events occur, *Taq* polymerases are unable to continue extending those particular DNA strands and so synthesis abruptly terminates (Innis et al. 1988; Barnes 1994). The production of these truncated PCR products, in turn, does not allow for a true "chain reaction" to take place because the number of templates for both primers does not exponentially increase. Thus PCRs using *Taq* as the sole DNA polymerase are incapable of generating a sufficient yield of accurately made PCR amplicons that are >5 kb in size.

A solution to the fidelity problem of *Taq* was proposed by Lundberg et al. (1991) who showed that, by replacing *Taq* in PCR with thermostable $3' \rightarrow 5'$ exonuclease (i.e., proofreading) DNA polymerases called *Pfu* (obtained from the thermophile **P***yrococcus* **fu***riosus*), they could improve the fidelity of final products by an order of magnitude. Because *Pfu* was evidently able to detect and excise its own nucleotide misincorporations, it might have seemed reasonable to believe that *Pfu*-driven PCRs might be a way to also solve the long PCR problem. Unfortunately, however, Barnes (1994) showed that such proofreading DNA polymerases are by themselves unable to sufficiently amplify loci longer than about 2 kb. Barnes hypothesized that the reason for these failures may be (oddly) due to the proofreading enzymes themselves. He reasoned that the addition of a sufficient amount of proofreading polymerase to amplify longer fragments may have the unfortunate side effect of rapid enzymatic degradation of the primers (i.e., exonucleases destroy single-stranded DNA). Therefore the presence of degraded primers in a running PCR may not only hinder the production of correct PCR products, but they can also lead to the production of PCR artifacts, which could adversely affect a DNA sequencing reaction. Interestingly, in the end, the use of such proofreading DNA polymerases proved to be a critical part of the solution to the long PCR problem.

The key was to use a *mixture* of two different types of thermostable DNA polymerases in order to take advantage of *Taq*'s high processivity along with the proofreading ability of $3' \rightarrow 5'$ exonuclease polymerases such as *Pfu*, Vent, or Deep Vent (Barnes 1994; Cheng et al. 1994a,b). The relative amount of each enzyme in the reaction was also important. The experiments by Barnes (1994) showed that by using a high level of *Taq* polymerase as a "parent" enzyme together with a low level (160–640-fold lower in terms of polymerase incorporation units) of a proofreading DNA polymerase, the *Taq*-caused nucleotide misincorporations could still be efficiently excised by the proofreading enzymes—thus allowing *Taq* to complete synthesis of those strands—but without suffering from excessive primer degradation caused by an abundance of proofreading polymerases. Using these enzyme mixes together with longer primers (27–33 bp) and modified thermocycling profiles, Barnes' approach provided an elegant resolution to the long PCR problem. This method, which is now commonly referred to as "long PCR" or "LA PCR" (for long and accurate), is capable of providing high product yields for loci as long as 35 or 40 kb (Barnes 1994; Cheng et al. 1994a,b).

In addition to the two-enzyme mix, the methodology of long PCR differs from standard PCR in several other ways. For example, a long PCR thermocycling profile can use the format:

Step 1: 94°C for 30 seconds (initial denaturation step)

Step 2: 94°C for 30 seconds

Step 3: 55–68°C for 30 seconds

Step 4: 68°C for 1 min/kb of target sequence

Repeat Steps 2–4 total of 30–35 times

You should notice several aspects of this profile that will differentiate it from a typical PCR profile. First, long PCR requires a significantly longer extension time. It is important to realize that *Taq* enzymes in commercially available long PCR "kits" have variable elongation rates (e.g., 1 kb per 15 seconds to 1 kb per minute) depending on the kit. Thus, extension times need to be calculated based on the kit specifications as well as the expected product length. Another difference is the primer annealing temperature. Long PCR tends to work best with PCR primers that are 30–33 bp (Barnes 1994; personal observation) and thus the optimal annealing temperatures for these longer primers will likely be higher than for more typical PCR primers. Some long PCR thermocycler profiles use the same temperature for both primer annealing and extension steps (e.g., Cheng et al. 1994a,b; Zardoya and Suárez 2008). Another difference between short and long PCR is the extension temperature. Notice that the extension temperature of 68°C is a little lower than the usual 70–72°C used in routine PCR. Although the optimal temperature for *Taq* polymerase is 70–74°C, Barnes (1994) found that extension temperatures above 68°C resulted in diminished long PCR product yields. Another factor limiting product yields is the duration of the denaturation step. Barnes (1994) conducted an experiment, which varied the time of the denaturation step and he observed that denaturing the DNA at 94°C for 20, 60, and 180 seconds resulted in progressively lower yields of an 8.4 kb product. He argued that the longer heat treatments stressed the DNA template enough to cause a higher incidence of depurination of sites, which can create problems because *Taq* polymerase stops synthesis at depurinated sites. Depurination of a site occurs when the glycosidic bond connecting the base with the sugar moiety is cleaved, which results in loss of the adenine or guanine base. Barnes also pointed out that a lower pH environment can worsen the depurination rate. Accordingly, based on these results as well as the study by Cheng et al. (1994a) it appears important to keep the denaturation step as brief as possible with the minimum temperature while also increasing the reaction pH to 8.8–9.2 (optimum of 9.1). Fortunately, a variety of long or LA PCR enzyme mixes with optimized buffers are commercially available so the researcher does not need to be concerned about fiddling with the pH of a long PCR. However, even with a proven commercially available long PCR reagent mix, there are additional "fine-tuning" procedures that the researcher can perform in order to improve the quality and yield of products. One variable is concentration of Mg^{2+} in the reaction mixture. As some product manuals will suggest, you should optimize the amount of free Mg^{2+} that is available for the polymerases. Secondly, improved specificity and target product yields can be achieved by hot starting the reactions (Cheng et al. 1994b; personal observations). This can be done by either purchasing a hot start long PCR kit or by performing a manual hot start.

5.7.3 Reverse Transcriptase-PCR

In phylogenomic studies concerned with the evolution of gene families it is essential to obtain DNA sequences for all homologous copies of a coding gene of interest that are present in a genome. However, because eukaryotic genes are usually split into different genomic segments consisting of exons and introns, this complicates the acquisition of full-length coding DNA sequences from genomic DNA templates. To address this challenge a variant of PCR called "RT-PCR" can be employed.

A summary of the RT-PCR method is as follows. The enzyme "RT" is used to synthesize DNA strands that are complementary to mature mRNA strands. As you will recall from Chapter 2, RT is a special type of DNA polymerase obtained from retroviruses that uses mRNA molecules as templates. Because most mature mRNAs have polyA tails, these stretches of sites can serve as annealing sites for polyT or "oligo(dT)" DNA primers, which, in turn, supplies a primer-template junction for the RT to begin synthesis of a DNA strand. These synthesized DNA strands are called *complementary DNAs* or "cDNAs." Normal PCR is then used to

amplify these cDNAs into double-stranded cDNAs. However, because DNA polymerases cannot synthesize DNA using the mRNA template strands, the mRNA strands must first be eliminated using the enzyme *RNase* H followed by a round of ssDNA synthesis using a DNA polymerase. But how does the DNA polymerase synthesize the strand complementary to the first cDNA strand without adding a primer? After removal of the mRNA strand, the 3′ end of the remaining single-stranded cDNA strand spontaneously forms a hairpin by looping back on itself to generate a short stretch of dsDNA. This hairpin structure creates the needed primer-template junction from which the DNA polymerase can synthesize a complementary strand resulting in fully double-stranded cDNAs. A special endonuclease called S1 nuclease, which excises single-stranded loops in DNA, is then used to cut the loop thereby creating double-stranded but separate cDNA molecules. This step completes the construction of a *cDNA library*, which can later be used with standard PCR to target and amplify specific genes using gene-specific primer pairs. The next major step is to use standard PCR to amplify these double-stranded cDNAs. Note that because the cDNA products of RT-PCR are derived from mature mRNA molecules they will not contain introns and thus they provide the uninterrupted gene sequence needed for phylogenomic analyses.

In the laboratory, the first step is to acquire high-quality (i.e., undegraded) RNA that has been purified. RNA is much more vulnerable to degradation than DNA via enzymatic destruction by *RNases* or chemical disintegration due to alkaline hydrolysis. Note that some DNA extraction kits include *RNase* A and therefore careful consideration should be given whether to destroy the RNA at the time of extraction because the RNA may be needed later for RT-PCR. Most standard PCRs will function fine in the presence of RNA and thus it may not be necessary to do the *RNase* treatment during a DNA extraction. *RNase* can be used later on the purified DNA in case the DNA is needed for an RNA-free application such as making some types of NGS libraries (Chapter 7). RT-PCR is done either in a "one-step" procedure in which the RT and PCRs are carried out in the same PCR tube or in a "two-step" process whereby the RT reaction is performed first in one tube then the reaction products are transferred to a second (or more) tubes for PCR amplification. Because the RT reagents are expensive compared to normal PCR reagents, the two-step method is good for the preparation of a cDNA library, which can then be aliquoted as templates for multiple different PCRs. It is critically important to take measures to avoid *RNase* contamination such as wearing gloves and using *RNase*-free filter tips. Regardless of whether a one- or two-step RT-PCR procedure is used, the reactions are driven by a thermocycler. The first cycle of an RT-PCR thermocycler program, which is specifically for the RT reaction, consists of three parts: (1) the temperature is set to about 50°C so that oligo(dT) primers can anneal to the polyA tails of the mRNAs; (2) in the next step, the temperature is lowered to the preferred temperature of the RT enzyme (37–42°C). This allows the RT to extend the cDNA strand along the RNA template starting at the primer-template junction; (3) in the last step of this first cycle, the temperature is raised to 92°C to inactivate the RT enzymes. The second cycle starts the PCR process using the newly formed double-stranded cDNAs as the DNA templates for amplification. At this time, forward and reverse gene-specific primers are needed to target particular cDNAs for amplification. The PCR uses a typical master mix containing *Taq* polymerase and runs for 30–40 cycles. After the reactions are completed, the success of the amplifications can be evaluated on a 1%–2% agarose gel. In the past, RT-PCR products would have been sequenced via laborious vector-cloning methods, but now NGS-based methods can sequence far larger numbers of such products in a more efficient manner.

REFERENCES

Barnes, W. M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc Natl Acad Sci USA* 91:2216–2220.

Brock, T. D. 1997. The value of basic research: Discovery of *Thermus aquaticus* and other extreme thermophiles. *Genetics* 146:1207–1210.

Brock, T. D. and H. Freeze. 1969. *Thermus aquaticus* gen. n. and sp. n., a non-sporulating extreme thermophile. *J Bacteriol* 98:289–297.

Brown, T. A. 2007. *Genomes* 3. New York: Garland Science/Taylor & Francis.

Cheng, S., S.-Y. Chang, P. Gravitt, and R. Respress. 1994b. Long PCR. *Nature* 369:684–685.

Cheng, S., C. Fockler, W. M. Barnes, and R. Higuchi. 1994a. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc Natl Acad Sci USA* 91:5695–5699.

Chou, Q., M. Russell, D. E. Birch, J. Raymond, and W. Bloch. 1992. Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucleic Acids Res* 20:1717–1723.

Innis, M. A., D. H. Gelfand, J. J. Sninsky, and T. J. White. 1990. *PCR Protocols: A Guide to Methods and Applications*. New York: Academic Press.

Innis, M. A., K. B. Myambo, D. H. Gelfand, and M. A. Brow. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc Natl Acad Sci USA* 85:9436–9440.

Kainz, P. 2000. The PCR plateau phase—Towards an understanding of its limitations. *Biochim Biophys Acta* 1494:23–27.

Keohavong, P. and W. G. Thilly. 1989. Fidelity of DNA polymerases in DNA amplification. *Proc Natl Acad Sci USA* 86:9253–9257.

Kleppe, K., E. Ohtsuka, R. Kleppe, I. Molineux, and H. G. Khorana. 1971. Studies on polynucleotides: XCVL repair replication of short synthetic DNA's as catalyzed by DNA polymerases. *J Mol Biol* 56:341–361.

Kornberg, A. and T. Baker. 1992. *DNA Replication*, 2nd edition. New York: W. H. Freeman and Company.

Korolev, S., M. Nayal, W. M. Barnes, E. Di Cera, and G. Waksman. 1995. Crystal structure of the large fragment of *Thermus aquaticus* DNA polymerase I at 2.5-A resolution: Structural basis for thermostability. *Proc Natl Acad Sci USA* 92:9264–9268.

Lundberg, K. S., D. D. Shoemaker, M. W. W. Adams, J. M. Short, J. A. Sorge, and E. J. Mathur. 1991. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* 108:1–6.

McPherson, M. J. and S. G. Moller. 2006. PCR, 2nd edition. Boca Raton: Taylor & Francis.

Mullis, K. B. and F. Faloona. 1987. Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. *Methods Enzymol* 155:335–350.

Mullis, K. B., F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich. 1986. Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* 11:263–273.

Palumbi, S. R. 1996. Chapter 7. Nucleic acids II: The polymerase chain reaction. In *Molecular Systematics*, 2nd edition, ed. D. M. Hillis, C. Moritz, and B. K. Mable, 205–247. Sunderland: Sinauer.

Saiki, R. K., D. H. Gelfand, S. Stoffel et al. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487–491.

Saiki, R. K., S. Scharf, F. Faloona et al. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230:1350–1354.

Sorenson, M. D. and T. W. Quinn. 1998. Numts: A challenge for avian systematics and population biology. *Auk* 115:214–221.

Tindall, K. R. and T. A. Kunkel. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* 27:6008–6013.

Watson, J. D., T. A. Baker, S. P. Bell, A. Gann, M. Levine, and R. Losick. 2014. *Molecular Biology of the Gene*, 7th edition. New York: Pearson Education, Inc.

Zardoya, R. and M. Suárez. 2008. Chapter 12. Sequencing and phylogenomic analysis of whole mitochondrial genomes in animals. In *Phylogenomics*, ed. W. J. Murphy, 185–200. Totowa: Humana Press.