Student Manual

Introduction to PCR — The Polymerase Chain Reaction

You are about to perform a procedure known as PCR[1] to amplify a specific sequence of your own DNA in a test tube. You will be looking for a particular piece of DNA that is present in the genes of many but not all people. Analysis of the data generated in this laboratory will enable you to determine whether your genomic DNA carries this piece of DNA, or **not**.

The genome, composed of DNA, is our hereditary code. This is the so-called "hard-wiring", the blueprint that controls much of why we look like we do, why we act like we do, and how we do the things we do. **Molecular biology** is the study of genes and the molecular details that regulate the flow of genetic information from DNA, to RNA and proteins, from generation to generation. **Biotechnology** uses this knowledge to manipulate organisms' (microbes, plants, or animals) DNA to help solve human problems.

Within the molecular framework of biology, DNA, RNA and proteins are closely tied to each other. Because proteins and enzymes ultimately play such a critical role in the life process, scientists have spent many lifetimes studying proteins in an attempt to understand how they work. With this understanding it was believed we could cure, prevent, and overcome disease and physical handicaps as well as explain exactly how and why organisms exist, propagate and die. However, the complete answer to how and why does not lie solely in the knowledge of how enzymes function; we must learn how they are made. If each enzyme is different, then what controls these differences and what is the blueprint for this difference? That answer lies within our genome, or genetic code.

Thus, you may realize why researchers today, in an attempt to understand the mechanisms behind the various biological processes, study nucleic acids as well as proteins in order to get a complete picture. In the last 20 years, many advances in the areas of nucleic acid techniques have allowed researchers the means to study the roles that nucleic acids play in biology. It took the imagination and hard work of many scientists to reveal the answers to one of the most mysterious puzzles of life — understanding the mechanisms that control how DNA is translated into proteins within living cells.

Before Beginning This Lab, See If You Can Answer the Following Questions

How is DNA faithfully passed on from generation to generation? What causes genetic disease in some people but not others? How do scientists obtain DNA to study? What secrets can DNA tell us about our origins? What human problems can an understanding of DNA help us solve? Should we unlock the secrets held in this most basic building block of life?

PCR Set the Stage for a Scientific Revolution

In 1983, Kary Mullis[2] at Cetus Corporation developed the molecular biology technique that has since revolutionized genetic research. This technique, called the **polymerase chain reaction** (PCR), transformed molecular biology into a multidisciplinary research field within 5 years of its invention. Before PCR, the molecular biology techniques used to study DNA required such a high level of expertise that relatively few scientists could use them.

The objective of PCR is to produce a large amount of DNA in a test tube (in vitro), starting from only a trace amount. Technically speaking, this means the controlled enzymatic amplification of a DNA sequence, or gene, of interest. The template strands can

be any form of double-stranded DNA such as genomic DNA. A researcher can take trace amounts of genomic DNA from a drop of blood, a single hair follicle, or a cheek cell (in theory, only a single template strand is needed to copy and generate millions of new identical DNA molecules) and make enough to study. Prior to PCR, this would have been impossible. It is the ability to amplify the precise sequence of DNA of interest that is the true power of PCR.

PCR has made an impact on four main areas of genetic research: gene mapping, cloning, DNA sequencing, and gene detection. PCR is now used as a medical diagnostic tool to detect specific mutations that may cause genetic disease[3], is used in criminal investigations and courts of law to identify suspects on the molecular level[4], and has been a powerful tool in the sequencing of the human genome[5]. Prior to PCR the use of molecular biology techniques for therapeutic, forensic, pharmaceutical, agricultural, or medical diagnostic purposes was neither practical nor cost-effective. The development of PCR technology transformed molecular biology from a difficult science to one of the most accessible and widely used disciplines of **biotechnology**.

Now, let's extract some of your own DNA.

Lesson 1 DNA Template Preparation

To obtain DNA for use in the polymerase chain reaction you will extract the DNA from your own living cells. It is interesting to note that DNA can be also extracted from mummies and fossilized dinosaur bones. In this lab activity, you will be isolating DNA from epithelial cells that line the inside of your cheek. This is accomplished by using a sterile pipet tip to gently scrape the inside of both your cheeks about 10 times each to scoop up the cells lining the surface. You will then boil the cells to rupture them and release the DNA they contain. To obtain pure DNA for PCR you will use the following procedure.

The cheek cells in the pipet tip are transferred into a micro test tube containing 200 µl of **InstaGene matrix**. This particulate matrix is made up of negatively charged microscopic beads that "chelate", or grab metal ions out of solution. It acts to trap metal ions, such as Mg^{2+}, which are required as catalysts or **cofactor**s in enzymatic reactions. Your cheek cells will then be lysed or ruptured by heating to release all of their cellular constituents, including enzymes that were once contained in the cheek cell lysosomes. Lysosomes are sacs within the cytoplasm that contain powerful enzymes, such as **DNases,** which are used by cells to digest the DNA of invading viruses. When you rupture the cells, these DNases can digest the released DNA of interest. However, when the cells are lysed in the presence of the chelating beads, the cofactors are adsorbed and are not available to the enzymes. This virtually blocks all enzyme degradation of the extracted DNA and results in a population of intact genomic DNA molecules that will be used as the template in your PCR reaction.

Your isolated cheek cells are first suspended in the InstaGene matrix and incubated at 56°C for 10 minutes. This "preincubation" step helps to soften the plasma membranes and release clumps of cells from each other. The increased temperature also acts to inactivate enzymes such as DNases, which will degrade the DNA template. After this 10 minute incubation period, the cells are then placed into a boiling (100°C) water bath for 6 minutes. The boiling ruptures the cells and releases the DNA from the cell nucleus. Your extracted genomic DNA will then be used as the target template for PCR amplification.

Lesson 1 DNA Template Preparation (Lab Protocol)

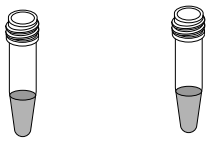
Workstation Checklist

Materials and supplies that should be present at the workstations prior to beginning this lab are listed below.

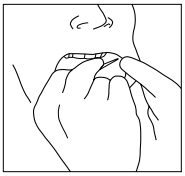
Student workstation	**Quantity per Station**	**(✔)**
Screwcap tubes with InstaGene™ matrix	8	❒
Foam micro test tube holder	2	❒
P-20 micropipet	1	❒
P-200 micropipet	1	❒
Pipet tips (filter type), 20–200 µl	12	❒
Permanent marker	1	❒
Copy of Quick Guide or protocol	1	❒
Waste container	1	❒
Instructor's workstation	**Quantity per Class**	
P-200 micropipet	1	❒
Pipet tips (filter type), 20–200 µl	1 box	❒
Pipet tips (filter type), 100–1,000 µl	1 box	❒
Water baths (56 and 100°C)	1 each	❒
Microcentrifuge	1	❒
or minicentrifuge	4	❒
Vortexer (optional)	1	❒

Lesson 1 DNA Template Preparation (Lab Protocol)

1. Each member of your team should have two screwcap tubes, each containing 200 µl of InstaGene matrix. Label the tube on the cap and on the side with your initials. In addition, label the tubes as "**tube 1**" and "**tube 2**". Each person should wash his or her hands before beginning step 2.

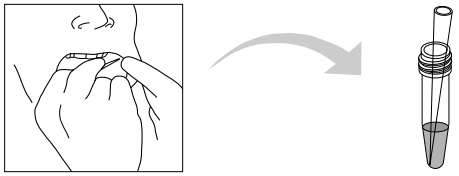


2. Using a sterile 20–200 µl filter pipet tip, gently scrape the insides of both cheeks 10 times each with the tip. This is most easily done by pinching and extending the corner of your mouth with one hand, and scraping the cheek with the tip in the other hand. Use firm but gentle pressure. The goal is to remove epithelial cells from the surface layer of your cheek lining. You should see a small volume of white cells in the pipet tip. Visually inspect the pipet tip to ensure that 0.5–1.0 mm of a cell plug is present.

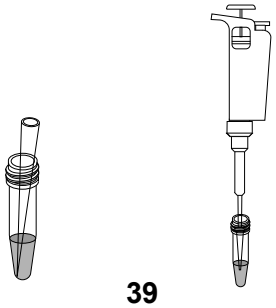


3. Place the tip that contains your cheek cells into the screwcap tube labeled "**tube 1**".

4. Using a second sterile 20–200 µl filter pipet tip, gently scrape the insides of both cheeks 10 times each with the tip. Place the tip that contains your cheek cells into your screwcap tube labeled as "**tube 2**".

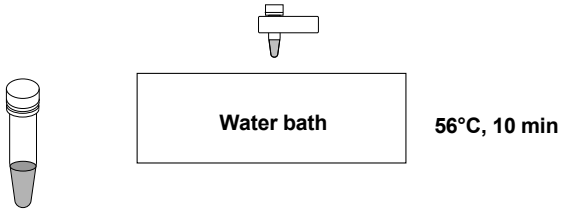


5. Place each tip on the end of a P-200 micropipet that is set on a 100 µl setting. Pipet up and down 5 times into the InstaGene matrix — the action of pipetting up and down mixes and transfers your cheek cells into the matrix.

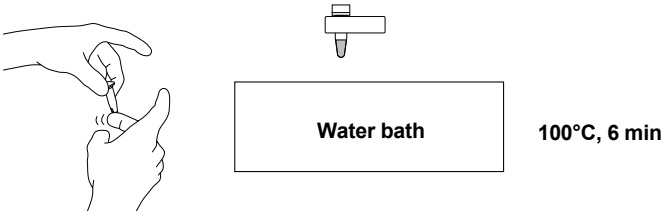


6. Screw the caps tightly on the tubes. Shake or vortex to mix the contents.

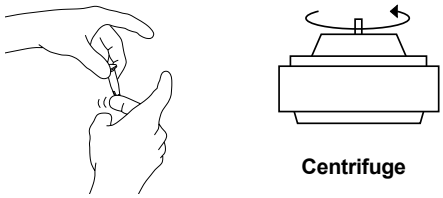
7. Place the tubes in the foam micro test tube holder. When all members of your team have collected their samples, float the holder with tubes in a 56°C water bath for 10 minutes. At the halfway point (5 minutes), remix the contents of the tubes by shaking or vortexing several times and place back in the water bath for the remaining 5 minutes.



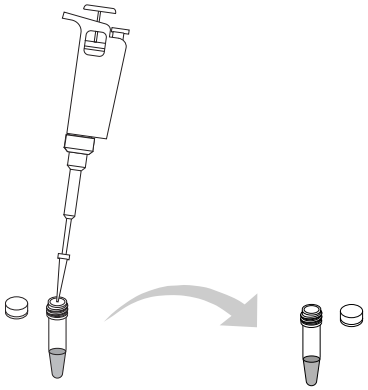
8. Remove the tubes from the water bath and remix by shaking the tubes several times. Now float the holder with tubes in a 100°C water bath for 6 minutes.



9. After 6 minutes, remove the tubes from the 100°C water bath and shake or vortex several times to resuspend the sample. Place the eight tubes in a balanced arrangement in a centrifuge. Pellet the matrix by spinning for 5 minutes at 6,000 x g (or 10 minutes at 2,000 x g) in the centrifuge.



10. Using a 200 µl pipet tip, remove 170 µl of the supernatant from your tube 1 and transfer the supernatant into tube 2. You now have one screwcap tube that contains your isolated genomic DNA.



11. Store your screwcap tube in the refrigerator until the next laboratory period or proceed to Step 2 of Lesson 2 if your teacher instructs you to do so.

Lesson 1 DNA Template Preparation

Focus Questions

1. Why is it necessary to trap the metal ions in the cheek cell solution before the boiling/lysis step at 100°C? What would happen if you did not put in the InstaGene matrix?

2. What is needed from the cheek cells in order to conduct the polymerase chain reaction?

3. What structures must be broken in order to release the DNA from a cell?

4. Why do you think the extracted cheek cell DNA is stored cold in the InstaGene matrix after boiling the samples?

Lesson 2 PCR Amplification

It is estimated that there are 30,000–50,000 individual genes in the human genome. The true power of PCR is the ability to target and make millions of copies of (or **amplify**) a specific piece of DNA (or gene) out of a complete genome. In this activity, you will amplify a region within your chromosome 16.

The recipe for a PCR amplification of DNA contains a simple mixture of ingredients. To replicate a piece of DNA, the reaction mixture requires the following components:

1. DNA template — containing the intact sequence of DNA to be amplified
2. Individual deoxynucleotides (A, T, G, and C) — raw material of DNA
3. DNA polymerase — an enzyme that assembles the nucleotides into a new DNA chain
4. Magnesium ions — a cofactor (catalyst) required by DNA polymerase to create the DNA chain
5. Oligonucleotide primers — pieces of DNA complementary to the template that tell DNA polymerase exactly where to start making copies
6. Salt buffer — provides the optimum ionic environment and pH for the PCR reaction

The template DNA in this exercise is genomic DNA that was extracted from your cheek cells. The complete master mix contains *Taq* DNA polymerase, deoxynucleotides, oligonucleotide primers, magnesium ions, and buffer. When all the other components are combined under the right conditions, a copy of the original double-stranded template DNA molecule is made — doubling the number of template strands. Each time this cycle is repeated, copies are made from copies and the number of template strands doubles — from 2 to 4 to 8 to 16 and so on — until after 20 cycles there are 1,048,554 exact copies of the target sequence.

PCR makes use of the same basic processes that cells use to duplicate their DNA.

1. Complementary DNA strand hybridization

2. DNA strand synthesis via DNA polymerase

The two DNA primers provided in this kit are designed to flank a DNA sequence within your genome and thus provide the exact start signal for the DNA polymerase to "zero in on" and begin synthesizing (replicating) copies of that target DNA. Complementary strand hybridization takes place when the two different **primers** anneal, or bind to each of their respective complementary base sequences on the template DNA.

The primers are two short single-stranded DNA molecules (23 bases long), one that is complementary to a portion of the 5'–3' strand, and another that is complementary to a portion of the 3'–5' strand of the template. These primers anneal to the separated template strands and serve as starting points for DNA *Taq* replication by DNA polymerase.

Taq DNA polymerase extends the annealed primers by "reading" the template strand and synthesizing the complementary sequence. In this way, *Taq* polymerase replicates the two template DNA strands. This polymerase has been isolated from a heat-stable bacterium (*Thermus aquaticus*) which in nature lives within the steam vents in Yellowstone National Park.[6] For this reason the enzymes within these bacteria have evolved to withstand high temperatures (94°C) and can be used in the PCR reaction.

PCR Step by Step

PCR amplification includes three main steps, a **denaturation step**, an **annealing step,** and an **extension step** (summarized in Figure 9). In denaturation the reaction mixture is heated to 94°C for 1 minute, which results in the melting or separation of the double-stranded DNA template into two single stranded molecules. In PCR amplification, DNA templates must be separated before the polymerase can generate a new copy. The high temperature required to melt the DNA strands normally would destroy the activity of most enzymes, but because *Taq* polymerase was isolated from bacteria that thrive in the high temperatures of hot springs, it remains active.

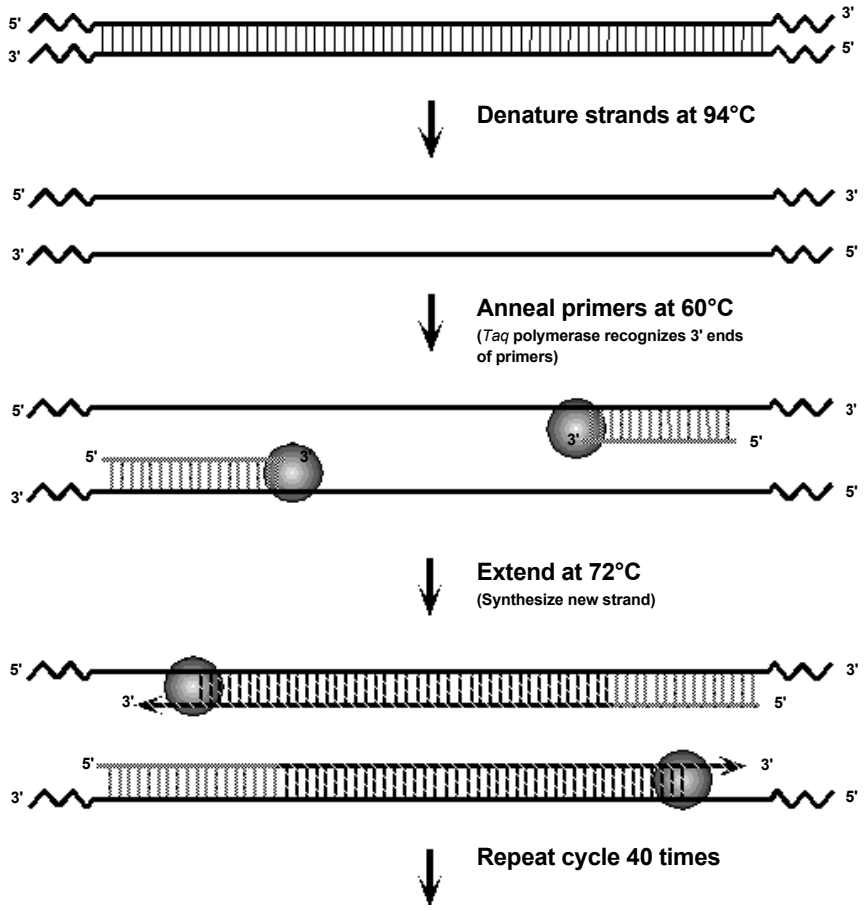


Fig. 9. A complete cycle of PCR.

During the annealing step, the oligonucleotide primers "anneal to" or find their complementary sequences on the two single-stranded template strands of DNA. In these annealed positions, they can act as primers for *Taq* DNA polymerase. They are called primers because they "prime" the synthesis of a new strand by providing a short sequence of double-stranded DNA for *Taq* polymerase to extend from and build a new complementary strand. Binding of the primers to their template sequences is also highly dependent on temperature. In this lab exercise, a 60°C annealing temperature is optimum for primer binding.

During the extension step, the job of *Taq* DNA polymerase is to add nucleotides (A, T, G, and C) one at a time to the primer to create a complementary copy of the DNA template.

During polymerization the reaction temperature is 72°C, the temperature that produces optimal *Taq* polymerase activity. The three steps of denaturation, annealing, and extension form one "cycle" of PCR, and a complete PCR amplification undergoes 40 cycles.

The entire 40 cycle reaction is carried out in a test tube placed into a thermal cycler. The thermal cycler contains an aluminum block that holds the samples and can be rapidly heated and cooled across broad temperature differences. The rapid heating and cooling of this thermal block is known as **temperature cycling** or **thermal cycling**.

Temperature Cycle = Denaturation Step (94°C) + Annealing Step (60°C) + Extension Step (72°C)

Lesson 2 PCR Amplification (Lab Protocol)

Workstation Checklist

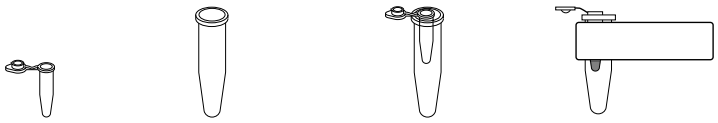
Materials and supplies that should be present at the workstations prior to beginning this lab are listed below.

Student workstation	**Quantity per Station**	**(✔)**
PCR tubes	4	❒
Micro test tubes, capless	4	❒
Complete master mix (with primers) on ice	1 tube	❒
P-20 micropipet	1	❒
Pipet tips (filter type), 2–20 µl	8	❒
Ice bucket with ice	1	❒
Foam micro test tube holders	2	❒
Permanent marker	1	❒
Waste container	1	❒
Copy of Quick Guide or protocol	1	❒
Instructor's (common) workstation	**Quantity per Class**	
Gel trays	1 per 2 stations	❒
Molten agarose	40 ml per gel	❒
Lab tape for gel trays	1 per station	❒
Gene Cycler or MyCycler thermal cycler	1	❒
Microcentrifuge	1	❒
or minicentrifuge	4	❒

Lesson 2 PCR Amplification (Lab Protocol)

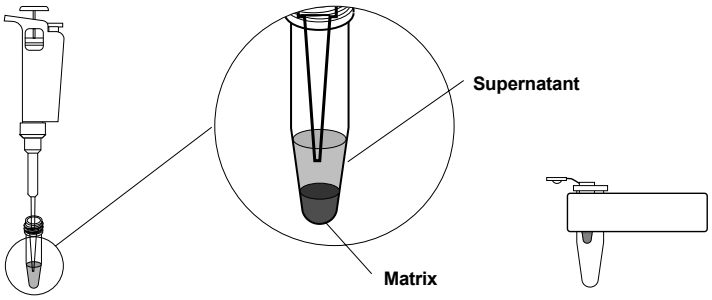
1. Obtain your screwcap tube that contains your genomic DNA template from the refrigerator and place on your lab bench. Centrifuge your tubes for 2 minutes at 6,000 x g or for 5 minutes at 2,000 x g in a centrifuge.

2. Each member of the team should obtain a PCR tube and capless micro test tube. Label each PCR tube on the side of the tube with your initials and place the PCR tube into the capless micro test tube as shown. Place the PCR tube in the foam micro test tube holder.

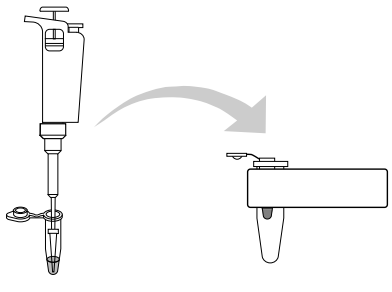


PCR tube Capless microtube

3. Transfer 20 µl of your DNA template from the supernatant in your screwcap tube into the bottom of the PCR tube. **Do not transfer any of the matrix beads into the PCR reaction because the reaction will be inhibited**.

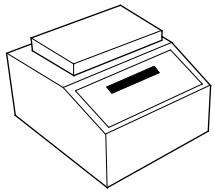


4. Locate the tube of yellow PCR master mix (labeled "Master") in your ice bucket. Transfer 20 µl of the master mix into your PCR tube. Mix by pipetting up and down 2–3 times. Cap the PCR tube tightly and keep on ice until instructed to proceed to the next step.



Master mix

5. Remove your PCR tube from the capless micro test tube and place the tube in the Gene Cycler or MyCycler.



6. When all of the PCR samples are in the thermal cycler, the teacher will begin the PCR reaction. The reaction will undergo 40 cycles of amplification, which will take approximately 3 hours.

7. If your teacher instructs you to do so, you will now pour your agarose gels (the gels may have been prepared ahead of time by the teacher).

Lesson 2 PCR Amplification

Focus Questions

1. Why is it necessary to have a primer on each side of the DNA segment to be amplified?

2. How did *Taq* polymerase acquire its name?

3. Why are there nucleotides (A, T, G, and C) in the master mix? What are the other components of the master mix, and what are their functions?

4. Describe the three main steps of each cycle of PCR amplification and what reactions occur at each temperature.

5. Explain why the precise length target DNA sequence doesn't get amplified until the third cycle. You may need to use additional paper and a drawing to explain your answer.

Lesson 3 Gel Electrophoresis of Amplified PCR Samples and Staining of Agarose Gels

What Are You Looking At?

Before you analyze your PCR products, let's take a look at the target sequence being explored.

What Can Genes and DNA Tell Us?

It is estimated that the 23 pairs, or 46 **chromosomes**, of the human genome (23 chromosomes come from the mother and the other 23 come from the father) contain approximately 30,000–50,000 genes. Each chromosome contains a series of specific genes. The larger chromosomes contain more DNA, and therefore more genes, compared to the smaller chromosomes. Each of the homologous chromosome pairs contains similar genes.

Each gene holds the code for a particular protein. Interestingly, the 30,000–50,000 genes only comprise 5% of the total chromosomal DNA. The other 95% is noncoding DNA. This noncoding DNA is interspersed in blocks between functional segments of genes and within genes, splitting them into segments. The exact function of the noncoding DNA is not known, although it is thought that noncoding DNA allows for the accumulation of mutations and variations in genomes.

When RNA is first transcribed from RNA, it contains both coding and noncoding sequences. While the RNA is still in the nucleus, the noncodong **introns** (in = stay with**in** the nucleus), are removed from the RNA while the **exons** (ex = **ex**it the nucleus) are spliced together to form the complete messenger RNA coding sequence for the protein (see Figure 10). This process is called **RNA splicing** and is carried out by specialized enzymes called **spliceosomes**.

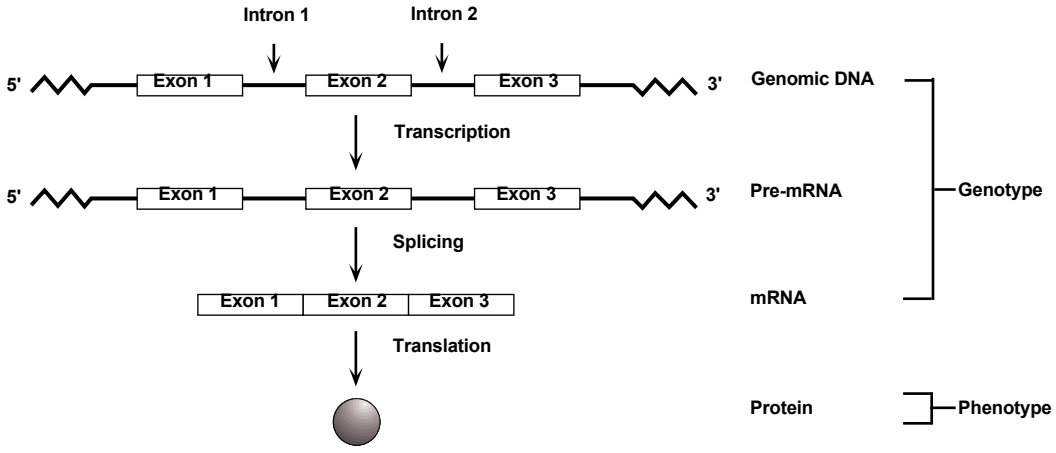


Fig. 10. Splicing of introns from genes.

Introns often vary in their size and sequence among individuals, while exons do not. This variation is thought to be the result of the differential accumulation of mutations in DNA throughout evolution. These mutations in our noncoding DNA are silently passed on to our descendants; we do not notice them because they do not affect our phenotypes. However, these differences in our DNA represent the molecular basis of DNA fingerprinting used in human identification and studies in population genetics.

The Target Sequence

The human genome contains small repetitive DNA elements or sequences that have become randomly inserted into it over millions of years. One such repetitive element is called the "Alu sequence"[7] (see Figure 11). This is a DNA sequence about 300 base pairs long that is repeated almost 500,000 times throughout the human genome.[8] The origin and function of these repeated sequences is not yet known.

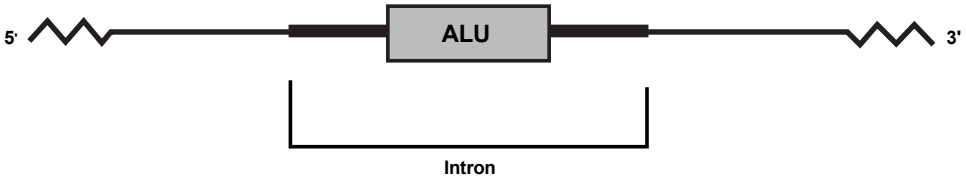


Fig. 11. Location of an Alu repetitive element within an intron.

Some of these Alu sequences have characteristics that make them very useful to geneticists. When present within introns of certain genes, they can either be associated with a disease or be used to estimate relatedness among individuals. In this exercise, analysis of a single Alu repeat is used to estimate its frequency in the population and as a simple measure of molecular genetic variation — **with no reference to disease or relatedness among individuals**.

In this laboratory activity you will look at an Alu element in the PV92 region of chromosome 16. This particular Alu element is **dimorphic,** meaning that the element is present in some individuals and not others. Some people have the insert in one copy of chromosome 16 (**one allele**), others may have the insert in both copies of chromosome 16 (**two alleles**), while some may not have the insert on either copy of the chromosome (see Figure 12). The presence or absence of this insert can be detected using PCR followed by agarose gel electrophoresis.

Since you are amplifying a region of DNA contained within an intron, the region of DNA is never really used in your body. So if you don't have it, don't worry.

The primers in this kit are designed to bracket a sequence within the PV92 region that is 641 base pairs long if the intron does not contain the Alu insertion, or 941 base pairs long if Alu is present. This increase in size is due to the 300 base pair sequence contributed by the Alu insert. When your PCR products are electrophoresed on an agarose gel, three distinct outcomes are possible.

If both chromosomes contain Alu inserts, each amplified PCR product will be 941 base pairs long. On a gel they will migrate at the same speed so there will be one band that corresponds to 941 base pairs. If neither chromosome contains the insert, each amplified PCR product will be 641 base pairs and they will migrate as one band that corresponds to 641 base pairs. If there is an Alu insert on one chromosome but not the other, there will be one PCR product of 641 base pairs and one of 941 base pairs. The gel will reveal two bands for such a sample.

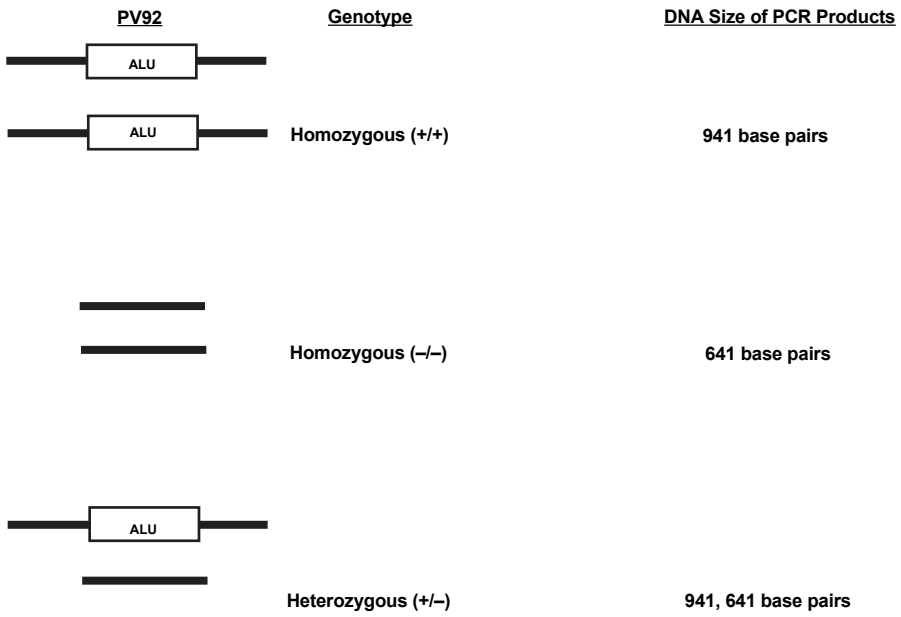


Fig. 12. The presence or absence of the Alu insert within the PV92 region of chromosome 16.

Electrophoresis separates DNA fragments according to their relative sizes (molecular weights). DNA fragments are loaded into an agarose gel slab, which is placed into a chamber filled with a conductive buffer solution. A direct current is passed between wire electrodes at each end of the chamber. DNA fragments are negatively charged, and when placed in an electric field will be drawn toward the positive pole and repelled by the negative pole. The matrix of the agarose gel acts as a molecular sieve through which smaller DNA fragments can move more easily than larger ones. Over a period of time, smaller fragments will travel farther than larger ones. Fragments of the same size stay together and migrate in what appears as a single "band" of DNA in the gel. In the sample gel below (Figure 13), PCR-amplified bands of 941 bp and 641 bp are separated based on their sizes.

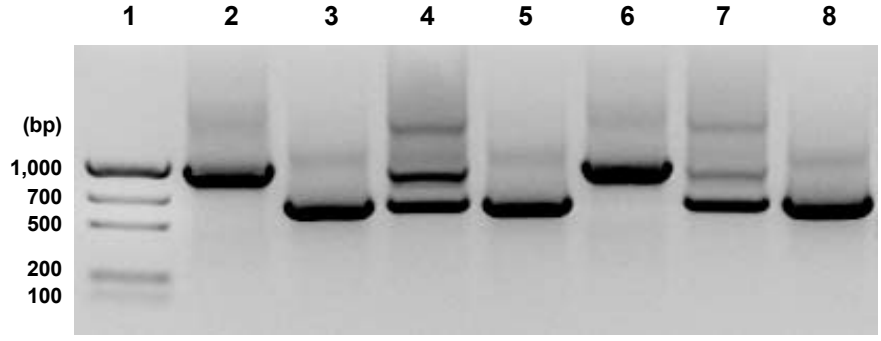


Fig. 13. Electrophoretic separation of DNA bands based on size. EZ Load DNA molecular mass ruler, which contains 1,000 bp, 700 bp, 500 bp, 200 bp, and 100 bp fragments (lane 1); two homozygous (+/+) individuals with 941 bp fragments (lanes 2, 6); three homozygous (–/–) individuals with 641 bp fragments (lanes 3, 5, and 8), and two heterozygous (+/–) individuals with 941/641 bp fragments (lanes 4 and 7).

Lesson 3 Gel Electrophoresis of Amplified PCR Samples and Staining of Agarose Gels (Lab Protocol)

Workstation Checklist

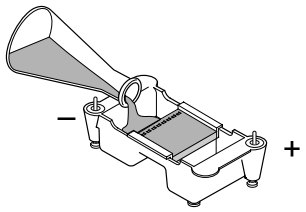
Materials and supplies that should be present at the workstations prior to beginning this lab are listed below.

Student workstation	**Quantity per Station**	**(✔)**
Agarose gel	1	❒
Student PCR samples	1/student	❒
MMR (DNA standard)	1 tube	❒
PV92 XC DNA loading dye	1 tube	❒
P-20 micropipet	1	❒
Pipet tips (filter type), 2–20 µl	12	❒
Permanent marker	1	❒
Foam micro test tube holder	1	❒
Gel box and power supply	1	❒
Fast Blast™ DNA stain, 1x or 100x solution	120 ml per 2 stations	❒
Gel support film (optional)	1	❒
Clear acetate sheets for tracing gels (optional)	1	❒
Warm tap water for destaining gels (if performing quick staining protocol)	1.5–2 L per 2 stations	❒
Large containers for destaining (if performing quick staining protocol)	1–3 per 2 stations	❒
Copy of Quick Guide or protocol	1	❒
Waste container	1	❒
Instructor's workstation	**Quantity per Class**	
1x TAE electrophoresis buffer	275 ml per gel box	❒
Amplified positive control samples (4 each)	12	❒
PV92 homozygous (+/+)		
PV92 homozygous (–/–)		
PV92 heterozygous (+/–)		
Shaking platform (optional)*	1	❒
Microcentrifuge	1	❒
or minicentrifuge	4	❒

*** Strongly recommended.**

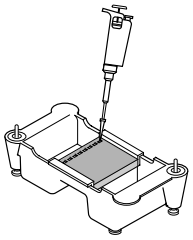
Lesson 3 Gel Electrophoresis of Amplified PCR Samples (Lab Protocol)

1. Remove your PCR samples from the thermal cycler and place in the micro test tube holder. If a centrifuge is available, place the PCR tubes in the capless micro test tubes and pulse-spin the tubes (~3 seconds at 2,000 x g) to bring the condensation that formed on the lids to the bottom of the tubes.
2. Add 10 µl of PV92 loading dye to each PCR tube and mix gently.
3. Obtain an agarose gel (either the one you poured or one pre-poured by your teacher). Place the casting tray with the solidified gel in it, onto the platform in the gel box. The wells should be at the cathode (–) end of the box, where the black lead is connected. Very carefully remove the comb from the gel by pulling it straight up, slowly.
4. Pour ~275 ml of electrophoresis buffer into the electrophoresis chamber, until it just covers the wells.

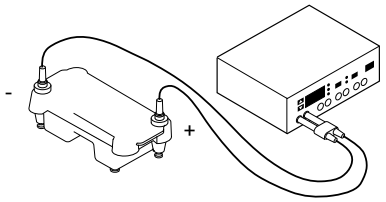


5. Using a clean tip for each sample, load the samples into the 8 wells of the gel in the following order:

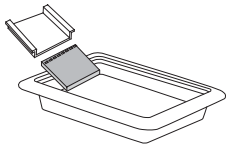
Lane	Sample	Load Volume
1	MMR (DNA standard)	10 µl
2	Homozygous (+/+) control	20 µl
3	Homozygous (–/–) control	20 µl
4	Heterozygous (+/–) control	20 µl
5	Student 1	20 µl
6	Student 2	20 µl
7	Student 3	20 µl
8	Student 4	20 µl



6. Secure the lid on the gel box. The lid will attach to the base in only one orientation: red to red and black to black. Connect the electrical leads to the power supply.
7. Turn on the power supply. Set it to 100 V and electrophorese the samples for 30 minutes.



8. When electrophoresis is complete, turn off the power and remove the lid from the gel box. Carefully remove the gel tray and the gel from the gel box. Be careful, the gel is very slippery. Nudge the gel off the gel tray with your thumb and carefully slide it into your plastic staining tray.



Staining of Agarose Gels

The moment of truth has arrived. What is your genotype? Are you a homozygote or a heterozygote? To find out, you will have to stain your agarose gel. Since DNA is naturally colorless, it is not immediately visible in the gel. Unaided visual examination of gel after electrophoresis indicates only the positions of the loading dyes and not the positions of the DNA fragments. DNA fragments are visualized by staining the gel with a blue dye called Fast Blast DNA stain. The blue dye molecules are positively charged and have a high affinity for the DNA. These blue dye molecules strongly bind to the DNA fragments and allow DNA to become visible. These visible bands of DNA may then be traced, photographed, sketched, or retained as a permanently dried gel for analysis.

Directions for Using Fast Blast DNA Stain

Below are two protocols for using Fast Blast DNA stain in the classroom. Use protocol 1 for quick staining of gels to visualize DNA bands in 12–15 minutes, and protocol 2 for overnight staining. Depending on the amount of time available, your teacher will decide which protocol to use. Two student teams will stain the gels per staining tray (you may want to notch gel corners for identification). Mark staining trays with initials and class period before beginning this activity.

WARNING
Although Fast Blast DNA stain is nontoxic and noncarcinogenic, latex or vinyl gloves should be worn while handling the stain or stained gels to keep hands from becoming stained blue. Lab coats or other protective clothing should be worn to avoid staining clothes.

Protocol 1: Quick Staining of Agarose Gels in 100x Fast Blast DNA Stain

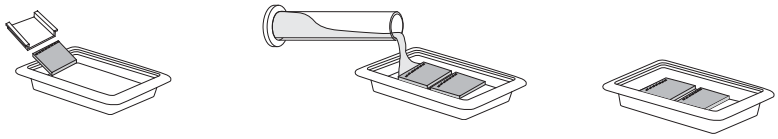
This protocol allows quick visualization of DNA bands in agarose gels within 15 minutes. For quick staining, Fast Blast DNA stain (500x) should be diluted to a 100x concentration. We recommend using 120 ml of 100x Fast Blast to stain two 7 x 7 cm or 7 x 10 cm agarose gels in individual staining trays provided in Bio-Rad's education kits. If alternative staining trays are used, add a sufficient volume of staining solution to completely submerge the gels.

Following electrophoresis, agarose gels must be removed from their gel trays before being placed in the staining solution. This is easily accomplished by holding the base of the gel tray in one hand and gently pushing out the gel with the thumb of the other hand. Because the gel is fragile, special attention must be given when handling it. We highly recommend using a large spatula or other supportive surface to transfer the gel from one container to another. Destaining requires the use of at least one large-volume container, capable of holding at least 500 ml, at each student workstation. Each student team may utilize separate washing containers for each wash step, or simply use a single container that is emptied after each wash and refilled for the next wash.

1. **Mark the staining tray with your initials and class period**. You will stain 2 gels per tray.

2. **Stain gels (2–3 minutes)**

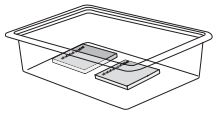
Remove each gel from the gel tray and carefully slide it into the staining tray. Pour approximately 120 ml of 100x stain into the staining tray. If necessary, add more 100x stain to completely submerge the gels. Stain the gels for 2–3 minutes, but not for more than 3 minutes. Using a funnel, pour the 100x stain into a storage bottle and save it for future use. **The stain can be reused at least 7 times**.



2–3 minutes

3. **Rinse gels**

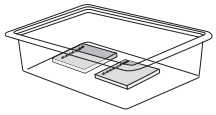
Transfer the gels into a large container containing 500–700 ml of clean, warm (40–55°C) tap water. Gently shake the gels in the water for ~10 seconds to rinse.



10 seconds

4. **Wash gels**

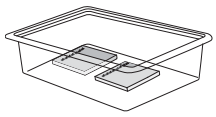
Transfer the gel into a large container with 500–700 ml of clean, warm tap water. Gently rock or shake the gels on a rocking platform for 5 minutes. If no rocking platform is available, move the gels gently in the water once every minute.



5 minutes

5. Wash gels

Perform a second wash as in step 4.



5 minutes

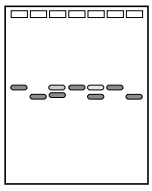
6. Record and Analyze results

Examine the stained gels for expected DNA bands. The bands may appear fuzzy immediately after the second wash, but will begin to develop into sharper bands within 5–15 minutes after the second wash. This is due to Fast Blast dye molecules migrating into the gel and binding more tightly to the DNA molecules.

To obtain maximum contrast, additional washes in warm water may be necessary. Destain to the desired level, but do not wash the gels in water overnight. If you cannot complete the destaining in the allocated time, you may transfer the gel to 1x Fast Blast stain for overnight staining. **See Protocol 2**.

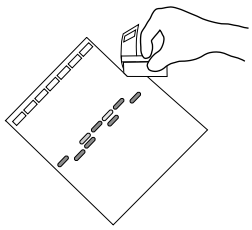
a. Place your gel on a light background and record your results by making a diagram as follows. Place a clear sheet of plastic sheet or acetate over the gel. With a permanent marker, trace the wells and band patterns onto the plastic sheet to make a replica picture of your gel. Remove the plastic sheet for later analysis. Alternatively, gels can be photocopied on a yellow piece of transparent film for optimal contrast.

b. With the help of your instructor, determine whether you are homozygous or heterozygous for the Alu insertion. First look at the control samples and note the migration patterns of the homozygous +/+, the homozygous –/–, and the heterozygous +/– samples (also refer to the example on page 51). You may notice that in the heterozygous sample the smaller 641 base pair band is more intense than the larger 941 bp band. This difference is due to the fact that the smaller fragment is amplified more efficiently than the larger fragment. Copies of the shorter fragment can be made at a faster rate than the bigger fragment, so more copies of the shorter fragment are created per cycle. Refer to pp. 61–64 for more information on how to analyze your data.

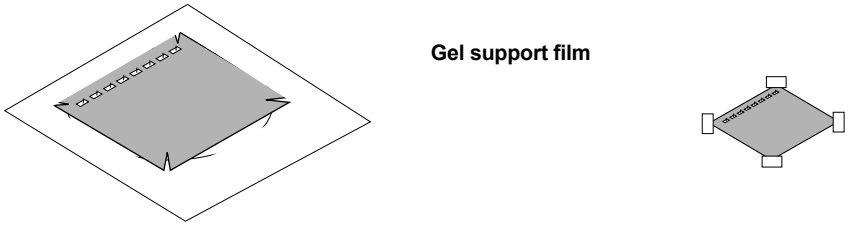


c. Dry the agarose gel as a permanent record of the experiment.

   i. Trim away any unloaded lanes with a knife or razor blade. Cut your gel from top to bottom to remove the lanes that you did not load samples into.



ii. Place the gel directly upon the hydrophilic size of a piece of gel support film. (Water will form beads on the hydrophobic side of a piece of gel support film.) Center the gel on the film and remove bubbles that may form between the gel and film. Place the film on a paper towel and let the gel dry in a well-ventilated area, making sure to avoid direct exposure to light. As the gel dries it will bond to the film but will not shrink. If left undisturbed on the support film, the gel will dry completely at room temperature after 2–3 days. The result will be a flat, transparent, and durable record for the experiment. Tape the dried gel into your laboratory notebook.



Note: Avoid extended exposure of dried gels to direct light to prevent band fading. However DNA bands will reappear if the dried gels are stored in the dark for 2–3 weeks after fading.

Protocol 2: Overnight Staining of Agarose Gels in 1x Fast Blast DNA Stain

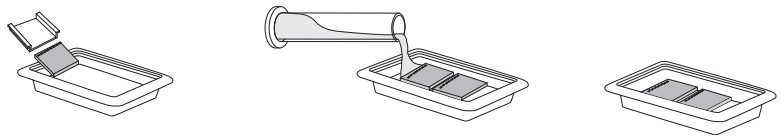
For overnight staining, Fast Blast DNA stain (500x) should be diluted to a 1x concentration. We recommend using 120 ml of 1x Fast Blast to stain two 7 x 7 cm or 7 x 10 cm agarose gels in individual staining trays provided in Bio-Rad's education kits. If alternative staining trays are used, add a sufficient volume of staining solution to completely submerge the gels.

Following DNA electrophoresis, agarose gels must be removed from their gel trays before being placed in the staining solution. This is easily accomplished by holding the base of the gel tray in one hand and gently pushing out the gel with the thumb of the other hand. Because the gel is fragile, special attention must be given when handling it.

1. Mark staining trays with your initials and class period. You will stain 2 gels per tray.

2. Stain gels (overnight)*

Pour 1x stain into a gel staining tray. Remove each gel from the gel tray and carefully slide it into the staining tray containing the stain. If necessary, add more 1x staining solution to completely submerge the gels. Place the staining tray on a rocking platform and agitate overnight. If no rocking platform is available, agitate the gels staining tray a few times during the staining period. You should begin to see DNA bands after 2 hours, but at least 8 hours of staining is recommended for complete visibility of stained bands.



Stain overnight

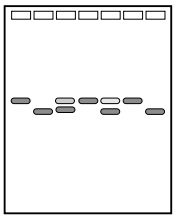
3. Analyze results

No destaining is required after staining with 1x Fast Blast. The gels can be analyzed immediately after staining.

a. Place your gel on a light background and record your results by making a diagram as follows. Place a clear sheet of plastic sheet or acetate over the gel. With a permanent marker, trace the wells and band patterns onto the plastic sheet to make a replica picture of your gel. Remove the plastic sheet for later analysis. Alternatively, gels can be photocopied on a yellow piece of transparent film for optimal contrast.

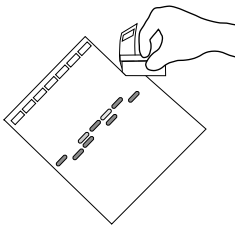
* Shake the gels gently and intermittently during overnight staining in 1x Fast Blast DNA stain; small DNA fragments tend to diffuse without shaking.

b. With the help of your instructor, determine whether you are homozygous or heterozygous for the Alu insertion. First look at the control samples and note the migration patterns of the homozygous +/+, the homozygous –/–, and the heterozygous +/– samples (also refer to the example on page 51). You may notice that in the heterozygous sample the smaller 641 base pair band is more intense than the larger 941 bp band. This difference is due to the fact that the smaller fragment is amplified more efficiently than the larger fragment. Copies of the shorter fragment can be made at a faster rate than the bigger fragment, so more copies of the shorter fragment are created per cycle.

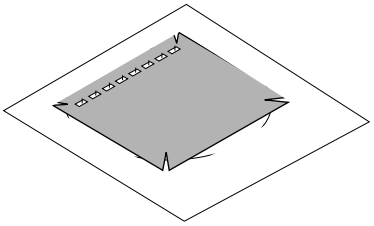


c. Dry the agarose gel as a permanent record of the experiment.

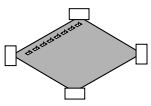
   i. Trim away any unloaded lanes with a knife or razor blade. Cut your gel from top to bottom to remove the lanes that you did not load samples into.



   ii. Place the gel directly upon the hydrophilic size of a piece of gel support film. (Water will form beads on the hydrophobic side of a piece of gel support film.) Center the gel on the film on a paper towel and let the gel dry in a well-ventilated area, making sure to avoid direct exposure to light. As the gel dries it will bond to the film but will not shrink. If left undisturbed on the support film, the gel will dry completely at room temperature after 2–3 days. The result will be a flat, transparent, and durable record for the experiment. Tape the dried gel into your laboratory notebook.



Gel support film



Note: Avoid extended exposure of dried gels to direct light to prevent band fading. However DNA bands will reappear if the dried gels are stored in the dark for 2–3 weeks after fading.

Lesson 3 Gel Electrophoresis of Amplified PCR Samples

Focus Questions

1. Explain the difference between an intron and an exon.

2. Why do the two possible PCR products differ in size by 300 base pairs?

3. Explain how agarose electrophoresis separates DNA fragments. Why does a smaller DNA fragment move faster than a larger one?

4. What kind of controls are run in this experiment? Why are they important? Could others be used?

Lesson 4 Analysis and Interpretation of Results

If the overnight staining protocol was used to stain the gels, record your results and dry gels as described earlier.

Analysis

Compare your sample lanes with the control lanes, using the DNA size marker as a reference. Mark the location and size of your fragment, or fragments. By comparing your DNA migration pattern to the controls, determine whether you are homozygous +/+, homozygous –/–, or heterozygous +/–. If your sample lane is blank, discuss with your classmates and teacher the possible reasons for lack of amplification.

Remember that the interpretation of this gel allows you to determine your genetic makeup only at the site, or **gene locus** (location), being studied. There are three possible genotypes for the Alu insert at the location you have amplified. For a class, determine the number of individuals of each genotype, homozygous +/+, homozygous –/–, and heterozygous +/–. Tally the class results in the table on page 62.

A major factor affecting the reliability of DNA fingerprinting evidence in forensics is population genetics and genetic statistics. In humans there are hundreds of loci or DNA segments, like Alu, that can be selected and used for fingerprinting analysis. Depending on demographic factors such as ethnicity or geographic isolation, some populations show much less variation in particular DNA segments than others. The degree of variation will affect the statistical odds of more than one individual having the same sequence. If 33% (1 out of three individuals) of a given population has the same fingerprinting pattern for a certain DNA segment, then little information will be obtained from using that segment alone to identify an individual.

When performing a DNA fingerprint to identify a suspect in a criminal case or paternity suit, what you want is not a 1 out of 3 (1/3) chance of a match in a population. What you want is more like 1 in 10 million ($1/10^7$) chance of a match. The frequency of a particular DNA pattern turning up in a population becomes extremely low when a series of DNA segments is selected and amplified, rather than just one segment alone. Amplifying multiple DNA segments from a single sample of genomic DNA can serve as a powerful tool to discriminate between individuals in a population. For DNA fingerprinting to be admissible as evidence in a court of law, it is necessary to selectively amplify and analyze 30 to 40 different DNA segments on multiple chromosomes. Therefore, in analyzing how incriminating the DNA evidence is, one needs to ask the question:

Statistically, how many people in a population may have the same DNA pattern as taken from a crime scene: 1 in 1,000,000? 1 in 10,000? Or 1 in 10?

In actuality, the Alu insert such as the one you have "fingerprinted" in this lab has been used to study the geographic migration patterns of different human populations over the course of human evolution.[8] The data from these studies have been published, and your class samples can be compared to the data collected from much larger populations.

Lesson 4 Analysis and Interpretation of Results

Focus Questions

Remember that this Alu sequence is inserted into a non-coding region of the PV92 locus on chromosome 16, and is not related to a particular disease, nor does it code for any protein sequence. It's just non-coding DNA that can be used as a tool to study human genotypic frequencies.

Because Alu repeats have become integrated into the general population at random, the Alu insert in chromosome 16 is very useful in the study of the gene frequencies in localized human populations. Theoretically, in some geographically isolated populations all individuals may be homozygous +/+, in others the individuals may all be homozygous –/– while in a ("melting-pot") population the three genotypes (polymorphisms) may exist in equilibrium. The results you obtain in this lab provide a real life opportunity to use the Hardy-Weinberg equation to examine and study genotypic and allelic frequencies of the Alu insert in your class.

The Hardy-Weinberg equation, $\mathbf{p^2 + 2pq + q^2 = 1}$, describes the frequencies of alleles in the gene pool of an entire population. In this case the entire population is your class:

p^2 = the frequency of an individual homozygous +/+ for the Alu insert

q^2 = the frequency of an individual homozygous –/– for the lack of an Alu insert

2pq = the frequency of an individual heterozygous +/– for the Alu insert

By determining frequencies of the Alu genotype within your class population, the allelic frequencies can also be calculated. Additionally, the genotypic frequencies of your class population can be compared to published results of larger population sizes.[9]

1. What is your genotype for the Alu insert in your PV92 region?

2. What are the observed (actual) genotypic frequencies of +/+, +/–, or –/– in your class population? Fill in the table below with your class data.

Table 1. Observed Class Genotypic Frequencies

Category	Number of genotypes	Frequency (# of genotypes/Total)
Homozygous (+/+)	p^2 =	
Heterozygous (+/–)	pq =	
Homozygous (–/–)	q^2 =	
	Total =	= 1

3. What is the frequency of each allele in your overall class sample? Fill in the table below with your class data. Remember, a class of 32 students (N), will have a total of 2(N) = 64 alleles.

Table 2. Calculated Allelic Frequencies for the Class

Category	Number	Class Allelic Frequency
total (+) alleles = p		p/Total =
total (–) alleles = q		q/Total =
	Total alleles =	= 1.00

Hint: The expected allelic frequencies can be calculated from the Hardy-Weinberg equation using your data from Table 1.

The number of p alleles is determined by adding the two(+) alleles of (p^2) plus the one(+) allele from (2pq). Mathematically, this can be written as: $p = 2(p^2) + (2pq)$

The number of q alleles is determined by adding the 2(–) alleles of (q^2) plus the one(–) allele from (2pq). Mathematically, this can be written as: $q = 2(q^2) + (2pq)$

The allelic frequency can be calculated by dividing the number of each allele by the total number of alleles.

$$p = \frac{\text{number of p alleles}}{\text{total number of alleles}}$$

$$q = \frac{\text{number of q alleles}}{\text{total number of alleles}}$$

4. The following table presents data from a USA-wide random population study.

Table 3. Sampling of USA Genotypic Frequencies for Alu

Category	Number of each genotype	Frequency
Homozygous (+/+) = p^2	p^2 = 2,422	0.2422
Heterozygous (+/–) = 2pq	pq = 5,528	0.5528
Homozygous (–/–) = q^2	q^2 = 2,050	0.2050
Number of Samples Studied	Total = 10,000	= 1.00

Now, using the data above, calculate the allelic frequencies for the USA data as you did for your class population in Table 2.

Table 4. Calculated Allelic Frequencies for USA

Category	Number	Frequency
total (+) alleles = p		p/Total =
total (–) alleles = q		q/Total =
	Total alleles =	= 1.00

5. How does your actual class data for allelic frequencies compare with that of the random sampling of the USA population? Would you expect them to match? What reasons can you think of to explain the differences or similarities?

Lesson 5 Analysis of Classroom Data Using Bioinformatics

Bioinformatics is a discipline that integrates mathematical, statistical, and computer tools to collect and process biological data. Bioinformatics has become an important tool in recent years for analyzing the extraordinarily large amount of biological information that is being generated by researchers around the world. In Lesson 5, you will perform a bioinformatics exercise to investigate the genotypic frequencies for the Alu polymorphism in your class population and compare them with the genotypic frequencies of other populations.

Following PCR amplification and electrophoresis of your samples, you will analyze your experimental data to determine your genotypes for the Alu insertion within the PV92 locus on chromosome 16. The classroom genotype data can then be entered into Allele Server of Cold Spring Harbor Laboratory's Dolan DNA Learning Center. Allele Server is a Web-based database that contains genotype data from a range of populations around the world as well as other classrooms and teacher training workshops. It also provides a collection of statistical analysis tools to examine the Alu insertion polymorphism at the population level. You can either analyze your classroom data as an individual population or compare your population with other populations in the database.

Once you enter classroom data into Allele Server, you can perform a Chi-square analysis to compare the Alu genotype frequencies within the class population with those predicted by the Hardy-Weinberg equation. The genotypic frequencies of the class population can also be compared with the genotypic frequencies of another population in the database, using the Chi-square analysis.

Using Allele Server

Note: The Dolan DNA Learning Center web site is continually updated. Some of the following information may change.

1. On your Web browser, go to **vector.cshl.org**
2. Log in to Allele Server using the username and password your instructor provides.
3. Once you have logged in, follow instructions provided in the pop-up window for using Allele Server. You may also open a new window and go to **dnalc.org/help/sad/topic_3.html** to get more detailed instructions. Follow the detailed instructions on how to populate the workspace, analyze groups, compare groups, and query the database.

Remember that as a registered user, you may store any groups that you loaded in your personal *Allele Server* database and analyze them at your convenience.

Appendix A
Review of Molecular Biology

This section provides an overview of terms and concepts with which students should be familiar in order to get the most out of this lab.

Genome — The sequence of DNA molecules within the nucleus which code for all proteins for a given species. Each segment of DNA which encodes a given protein is called a gene. The information contained in the genome comprises the organism's hereditary code.

Molecular Biology — The study of genes and the molecular details which regulate the flow of genetic information from DNA, to RNA and proteins, and from generation to generation.

Biotechnology — The manipulation of organisms (microbes, plants or animals) DNA to help solve human problems.

Any living organism functions based on the complicated interactions between nucleic acids, proteins, lipids (fat), and carbohydrates. In nearly all cases, certain proteins termed enzymes control the almost infinite number of interactions and life processes in living creatures. Think of enzymes and proteins as all the different people on earth. Each person performs a different role, function, or job on this planet and although people are not the actual physical make-up of buildings, documents, food, and roads it is the people that make these buildings and roads and write the documents and plant and nurture the crops. In the same way, enzymes and proteins do not comprise bones, lipids, sex hormones, and sugars, but enzymes control these structures, their interactions, and processes.

Because proteins and enzymes ultimately play such a critical role in the life process, scientists have spent many lifetimes studying proteins in an attempt to understand how they work and how they can be controlled. With a complete understanding we could cure, prevent, and overcome many diseases and physical handicaps as well as explain exactly how and why organisms exist, propagate, and die. However, the complete answers do not lie solely in the knowledge of how enzymes function, we must learn how they are made. Before we can control enzymes, we must understand where they come from and what is the basis of the molecular information that encodes proteins. That answer lies within our **genetic code**.

Each living organism has its own blueprint for life. This blueprint defines how an organism will look and function (using enzymes as a means to form the appearance and control the functions). The blueprint codes for all the different enzymes. With amazing precision this blueprint gets passed on from generation to generation of each species for as long as that species continues to propagate.

The transfer of this blueprint from generation to generation is called **heredity**. The blueprint for any organism is called its **genome**. The hereditary code is encrypted within the sequence of the DNA molecules that make up the genome. The molecule that comprises the genome and thus the hereditary code is DNA (**deoxyribonucleic acid**).

The genome consists of very long DNA/protein complexes called **chromosomes**. Prokaryotes, organisms lacking a true nucleus, have only one chromosome. All other species, eukaryotes, have a defined cell nucleus which contains multiple chromosomes. The nucleus is a defined, membrane-enclosed region of the cell that contains the chromosomes. The number of chromosomes varies with the organism — from 2 or 3 in some yeasts and up to 100 or so in some fish. Humans have 46.

In most cases chromosomes come in nearly identical pairs (one member of the chromosome pair from each parent). In general, the members of a pair differ in small details from each other, since they come from different parents, but are otherwise identical or **homologous**. Cells with homologous pairs of chromosomes are called diploid. Nearly all cells of an organism are diploid. Cells that have only one chromosome of each pair are called haploid. All sperm and ova are haploid.

The process of forming sperm and ova is called **meiosis**. **Meiosis** starts with a diploid cell that divides into two haploid cells. When the sperm fertilizes the ovum the two nuclei fuse and thus the new nucleus contains pairs of each chromosome, one partner from each parent. The result is called a diploid zygote.

All cells of diploid organisms duplicate chromosomal pairs when they divide (except when sperm and ova are formed) so that all body cells (called somatic cells) of an organism are diploid. The process of cell division in which the chromosomes are duplicated and each daughter cell gets pairs of chromosomes is called **mitosis**. It is through the processes of mitosis and meiosis that the hereditary code is passed from cell to cell and generation to generation. Now that we know where the code is and how that code is passed on, we need to know how the code produces the enzymes that control life. The actual DNA code for a protein is contained within a segment of a chromosome called a gene. In nearly all cases, diploid organisms will have the same gene on a specific chromosome pair. The gene on each chromosome of a specific chromosome pair is also called an **allele**.

To clarify, a gene encodes a particular protein that performs a particular function. An allele refers to the same gene existing on each chromosome of a specific chromosome pair. Thus, there are genes for hair color and there is an allele for the hair color gene on each chromosome pair. The gene or allele DNA code can also be called the genotype.

When the protein is made from this code and performs its function, the physical trait or result that is seen is called the **phenotype**. In many cases the two alleles on the specific chromosome pair coding for a protein differ slightly in their respective DNA code (genotype). Any slight difference in code between the two alleles can result in two different proteins that although intended to perform basically the same function, actually carry out that function slightly differently, causing different results and thus different phenotypes.

Therefore, it is not only the various combinations of chromosomes a parent contributes to each offspring but also the various combinations of alleles and how each of the enzymes coded from the alleles work together that decide how we look and allow us to function. The various combinations are nearly infinite and that is why we are all different. The study of genotypes and phenotypes is often referred to as **Mendelian genetics** (after Mendel, the individual who pioneered the study of heredity and genetics).

DNA: What is it?

A DNA molecule is a long polymer consisting of four different components called **bases**. The four bases are also called **nucleotides**. It is the various combinations of these four bases or nucleotides that create a unique DNA code or sequence (also genotype, gene, and allele). Nucleotides are comprised of three different components:

- **nitrogen base**
- **deoxyribose sugar**
- **phosphate group**

Each nucleotide contains the same ribose sugar and the phosphate group. What makes each nucleotide unique is its nitrogen base. There are four nitrogen bases:

Adenine (A)

Thymine (T)

Guanine (G)

Cytosine (C)

A DNA nucleotide chain is created by the connection of the phosphate group to the ribose sugar of the next nucleotide. This connection creates the "backbone" of the DNA molecule.

To designate the different ends of this single-stranded chain we must make use of some typical biochemistry terminology. The carbons on any sugar are numbered. The ribose sugar of a nucleotide contains 5 carbons. The phosphate group (PO_4) of a given nucleotide is connected to the 5' carbon of its ribose sugar. This phosphate group is connected to the 3' carbon on the ribose sugar of the next nucleotide.

Thus, the end of a single-strand DNA molecule that has a free phosphate group (i.e., not attached to another nucleotide's ribose sugar) is called the 5' end, and the end of the single-strand chain that has an open hydroxyl group (OH) on the 3' carbon of the ribose sugar (because no subsequent phosphate of another nucleotide is attached) is called the 3' end (see Figure 14).

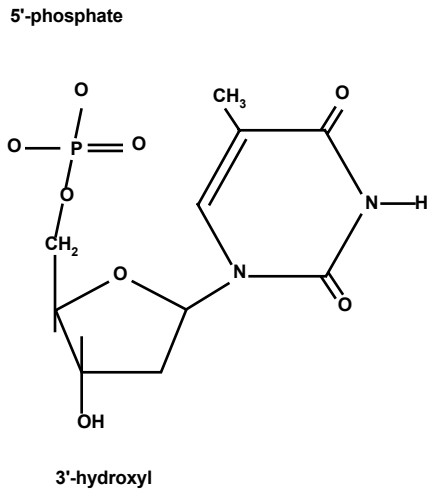


Fig. 14. Structure of one nucleotide of deoxyribonucleic acid.

It has become standard that single-stranded DNA molecule is written with the 5' end on the left and the 3' end on the right. Therefore, a single-stranded DNA chain's sequence is aligned from left to right starting on the left with the 5' nucleotide and moving to the right until the 3' nucleotide is last. Most DNA sequences are read 5' to 3'.

However, the long DNA molecules or chains that comprise the chromosomes are not single-stranded molecules. From X-ray crystallography patterns of DNA, and some imaginative molecular model building, Watson and Crick (along with some help from Messelson and Stahl) deduced that DNA is in fact a **double-stranded** molecule with the two single-strands of DNA held together by **hydrogen bonds** between the nitrogen bases

(A, T, G, and C). This double-stranded molecule is often called a duplex (see Figure 15). There are several important properties of double-stranded DNA molecules.

- Chromosomal (also called genomic) DNA is double-stranded.
- The overall structure is that of a helix with two strands intertwined.
- The structure can be viewed as a twisted ladder.
- The ribose-phosphate backbones are the sides of the ladder.
- The nitrogen bases (A, T, G, and C) hydrogen bonded to each other are the rungs.
- Only the nitrogen bases A and T and C and G can form hydrogen bonds to each other. When A binds to T or C binds to G this is considered **base pairing.** C and T, and A and G do not form hydrogen bonds.
- The orientation of the two strands is anti-parallel; that is, construction of the strands is in opposite directions. This means that the ladder runs 5' to 3' in a left to right direction for one strand and 3' to 5' in the left to right direction for the other strand.

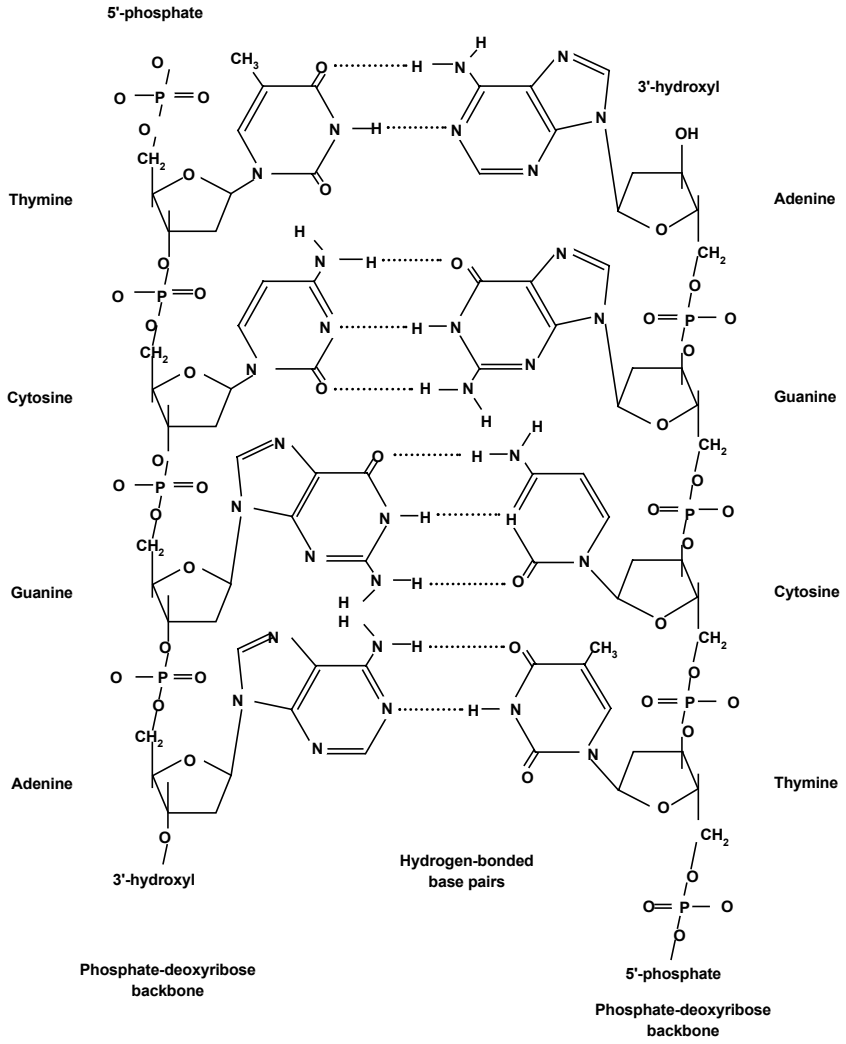


Fig. 15. Molecular structure of a portion of a double-stranded DNA molecule.

It follows From This Structure That

- Because A only binds to T and G only binds to C, the two strands will have exactly the opposite, or complementary, sequence running in opposite directions (one strand 5' to 3', the other 3' to 5').
- One strand is called the "sense" strand and the other is "antisense".
- These two complementary strands anneal or hybridize to each other through hydrogen bonds between the bases.
- A new single strand can be synthesized using its complement as the template for new strand synthesis.
- Each strand, then, carries the potential to deliver and code for the same information.

The length of any double-stranded DNA molecule is given in terms of base pairs (bp). If a DNA strand contains over a thousand base pairs, the unit of measure is kilobases (kb). If there are over one million base pairs in a strand the unit of measure is megabases (Mb).

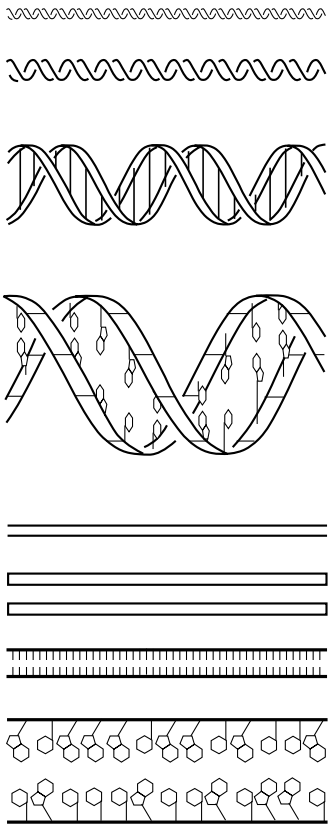


Fig. 16. DNA (deoxyribonucleic acid) — A long chainlike molecule that stores genetic information. DNA is graphically represented in a number of different ways, depending on the amount of detail desired.

DNA Replication — Strand Synthesis

New strands are synthesized by enzymes called **DNA polymerases**. New strands are always synthesized in a 5' to 3' direction. For a new single strand of DNA to be synthesized, another single strand is necessary. The single strand of DNA that will be used to synthesize its complementary strand is called the **template strand**.

However, in order for the DNA polymerase to start synthesizing a new complementary strand, a section of double-stranded DNA must be present for the DNA polymerase to initiate synthesis. A short strand (20–50 base pairs) used to anneal to the template strand to create the double-stranded start site for the DNA polymerase is called an oligonucleotide primer. This primer is almost always a short strand of nucleotides complementary to a part of the template where the researcher wants synthesis to begin. It must have a free 3' hydroxyl group (OH) for the DNA polymerase to attach the 5' phosphate of the next nucleotide.

The DNA polymerase grabs free (single) nucleotides from the surrounding environment and joins the 5' phosphate of the new nucleotide to the 3' hydroxyl group (OH) of the new complementary strand. This 5' to 3' joining process creates the backbone of the new DNA strand.

The newly synthesized strand maintains its complementarity with the template strand because the DNA polymerase only joins two nucleotides during new strand synthesis if the new nucleotide has its complement on the template strand. For example, the DNA polymerase will only join a G to the 3' end of the newly synthesized strand if there is the C counterpart on the template strand to form a hydrogen bond. Guanine will not be joined to the new strand if adenine, thymine or guanine is the opposite nucleotide on the template strand.

DNA polymerase and strand synthesis allow DNA to replicate during mitosis. Interestingly, both new DNA strands are synthesized simultaneously from the two original DNA template strands during mitotic DNA replication.

As you can see DNA, RNA, and proteins are closely tied to each other. Thus you can realize why researchers today, in an attempt to understand the mechanisms behind the various life processes, must study the nucleic acids as well as the proteins to get complete answers about the flow of information carried in the genetic code. In the last 20 years, many gains in the areas of nucleic acid techniques have finally allowed researchers the means to study the roles of nucleic acids in life processes.

Individual discoveries by many scientists have contributed the pieces that have begun to solve one of the most mysterious puzzles of life — understanding the hereditary code. In 1985, enough pieces of the puzzle were in place for a major breakthrough to occur. This elucidation of how the necessary molecular components interact to faithfully replicate DNA within living cells led to the development of a technique for creating DNA in a test tube. This technique is called the **polymerase chain reaction,** or **PCR**.

Appendix B
Glossary of Terms

Aliquot	The division of a quantity of material into smaller, equal parts.
Alu	A small piece of repetitive DNA that contains the *Alu* I restriction enzyme site, from which the sequence obtained its name.
Annealing	Binding of oligonucleotide primers to complementary sequences on the template DNA strands.
Chelation	To bind metal ions in solution. An example of a common chelating agent is EDTA or EthyleneDiamine Tetraacetic Acid.
Chelex®	Microscopic beads which chelate divalent cations and compose the InstaGene matrix.
Cofactors	Ions or small molecules needed by an enzyme to function properly. For example, *Taq* DNA polymerase needs Mg^{2+} in order to function properly. Mg^{2+} would therefore be considered a cofactor.
Denature	The process of melting apart two complementary DNA strands. In vivo denaturation is accomplished by enzymes; in the PCR reaction denaturation is accomplished by heat.
DNases	Digestive enzymes which degrade DNA.
dNTPs	Commonly used abbreviation for all four deoxynucleotide triphosphates (dATP, dTTP, dGTP, dCTP) used in synthesizing DNA.
Ethidium bromide	A fluorescent dye molecule which intercalates between DNA base pairs and fluoresces when exposed to ultraviolet light.
Eukaryotes	Organisms that are made up of cells containing a membrane-bound nucleus that contains the genetic material (DNA).
Exon	The region of a transcribed messenger RNA molecule that gets spliced together and leaves the nucleus for translation into protein sequence.
Extension	This refers to the process of *Taq* polymerase adding dNTPs (deoxynucleotide triphosphates — dATP, dTTP, dCTP, or dGTP) onto the ends of the oligonucleotide primers. Extension follows the base pairing rule and proceeds in the 5' to 3' direction.
Genomic DNA	The sum total of the DNA that is found within the nucleus of a cell.
Hardy-Weinberg equilibrium	The conditions that enable a population to maintain its genetic frequencies. These conditions are: large population, random mating, no immigration or emigration, no mutations, and no natural selection.

Homologous chromosomes	A pair of complementary chromosomes that contain the same genetic sequences, or genes, with one chromosome inherited from the mother and one chromosome inherited from the father.
InstaGene™ matrix	Microscopic beads that bind divalent cations in solution. The binding or sequestering of divalent cations prevents their availability to enzymes that can degrade the DNA template.
Intron	The region of a transcribed messenger RNA that is spliced out of the mRNA and is not translated into protein sequence.
Lysis	The process of rupturing a cell to release its constituents. In this exercise, human cheek cells are lysed to release genomic DNA for PCR reactions.
Master mix	The main solution of a PCR reaction which contains all of the necessary components (dNTPs, primer, buffer, salts, polymerase, magnesium) of the reaction except the template DNA.
Mendelian inheritance	For each inherited characteristic, an organism contains two genes, or alleles, one inherited from each parent. There can be two forms of alleles, dominant and recessive, and dominant alleles mask the expression of recessive alleles.
Messenger RNA	A type of RNA that is synthesized from the genetic material (DNA) and that attaches to ribosomes and is translated into protein.
Nucleotides	The fundamental unit of DNA or RNA. They consist of a sugar (deoxyribose or ribose), phosphate, and nitrogenous base (adenine, thymine, cytosine, or guanine and uracil in place of thymine in RNA).
PCR	Polymerase chain reaction. The process of amplifying or synthesizing DNA within a test tube.
Primers	A small series of nucleotides (usually 16–24 bases in length) that recognize a particular sequence of nucleotides on the target DNA sequence. Primers for the polymerase chain reaction are usually synthesized in a laboratory.
Reagents	Materials needed to conduct an experiment. They are usually solutions or mixtures of various solutions.
***Taq* DNA polymerase**	Heat stable DNA polymerase that was isolated from the heat stable bacteria *Thermus aquaticus*. This DNA polymerase is commonly used in PCR reactions.
Template	The strand of DNA that contains the target sequences of the oligonucleotide primers and that will be copied into its complementary strand.

Appendix C
PCR Amplification and Sterile Technique

PCR is a powerful and sensitive technique that enables researchers to produce large quantities of DNA from very small amounts of starting material. Because of this sensitivity, contamination of PCR reactions with unwanted, extraneous DNA is always a possibility. Therefore, utmost care must be taken to prevent cross-contamination of samples. Steps to be taken to prevent contamination and failed experiments include:

1. **Filter-type pipet tips**. The end of the barrel of micropipets can easily become contaminated with DNA molecules that are aerosolized. Pipet tips which contain a filter at the end can prevent aerosol contamination from micropipets. DNA molecules which are found within the micropipet can not pass through the filter and can not contaminate PCR reactions. Xcluda® aerosol barrier pipet tips (catalog #211-2006EDU and 211-2016EDU) are ideal pipet tips to use in PCR reactions.

2. **Aliquot reagents**. Sharing of reagents and multiple pipettings into the same reagent tube will most likely introduce contaminants into your PCR reactions. When at all possible, aliquot reagents into small portions for each team, or if possible, for each student. If an aliquotted reagent tube does become contaminated, then only a minimal number of PCR reactions will become contaminated and fail.

3. **Change pipet tips**. Always change pipet tips when entering a reagent for the first time. If a pipet tip is used repeatedly, contaminating DNA molecules on the outside of the tip will be passed into other solutions, resulting in contaminated PCR reactions. If you are at all unsure if your pipet tip is clean, use the safe rule of thumb and discard the tip and get a new one. The price of a few extra tips is a lot smaller than the price of failed reactions.

4. **Use good sterile technique**. When opening, aliquotting, or pipetting reagents, leave the tube open for as little time as possible. Tubes which are open and exposed to the air can easily become contaminated by DNA molecules that are aerosolized or which are present in your mouth/breath, etc. Go into reagent tubes efficiently, and close them when you are finished pipetting. Also, try not to pick tubes up by the rim or cap as you can easily introduce contaminating DNA molecules from your fingertips.