Amplifying the ALU intron for Hardy-Weinberg Analysis

OBJECTIVES:
- Review the following terms and concepts presented in Biology 201: enzymes, DNA structure and replication, role of DNA polymerase, primers, cell structure, PCR, gel electrophoresis, homologous chromosomes, introns, Mendelian Genetics, alleles
- Learn the technique of Polymerase Chain Reaction (PCR)
- Successfully extract DNA from cheek cells
- Successfully amplify the ALU intron
- Analyze the class data to assess Hardy-Weinberg Equilibrium
- Introduce population genetics

INTRODUCTION
In this Lab exercise, we will attempt to isolate our own DNA and then use the Polymerase Chain Reaction (PCR) to analyze our own genetic make-up! Recall that PCR is a powerful technique that mimics cellular DNA replication to make millions of copies of short, specific regions of DNA. We will use this technique to amplify a short, specific region of our own genomic DNA. Amplifying this DNA will allow us to compare the frequencies of this allele in our class (population) and compare it to those predicted by the Hardy-Weinberg Equation.

A Quick Review of PCR:
In 1983, Kary Mullis at Cetus Corporation developed the molecular biology technique known as the polymerase chain reaction (PCR). PCR revolutionized genetic research, allowing scientists to easily amplify short specific regions of DNA for a variety of purposes including gene mapping, cloning, DNA sequencing and gene detection.

The objective of PCR is to produce a large amount of DNA in a test tube (in vitro) starting from only a trace amount. This template 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 and make enough to study. Prior to PCR, this would have been impossible!

PCR is conducted in three steps: 1) Denature the template DNA, 2) Allow the primers to anneal, and 3) Extend (copy) the template DNA. In the first step, the template DNA is heated up to break the hydrogen bonds holding the two strands together. This allows each strand to serve as a template for generating copies of the DNA. In the second step, the temperature is reduced to allow the primes to anneal, or bind, at their complimentary sequence on the template. In the third step, the temperature is raised again to allow the DNA polymerase to bind at the primer and add nucleotides at its 3’ end. These steps are repeated 20-40 times, generating literally millions of copies of DNA from each template molecule!

The Target of Our PCR:
Recall that we humans have 23 pairs of chromosomes, or a total of 46 chromosomes. These chromosomes contain somewhere between 30,000 and 50,000 genes. Interestingly, these genes occupy only ~5% of our DNA. The other 95% of our DNA consists of non-coding DNA, or DNA that doesn’t directly code for a particular protein. Introns are an
example of some of this non-coding DNA. As you might imagine, the sequence of our introns is much more variable than the sequence of the coding portions of our genes (exons) as mutations in introns don’t necessarily impact the structure of the protein. We carry, and pass along, such mutations silently as they don’t affect our phenotype.

Over time, many short inserted sequences have accumulated in our DNA. For example, most of us carry approximately 500,000 copies of a 300 bp sequence known as the Alu sequence in our DNA. The origin and function of these sequences are still unknown. Despite this, these repeated Alu sequences have proved interesting for geneticists as when they are present in particular genes, they can be associated with diseases. The pattern of insertions can also be used to study the relatedness of individuals.

The target of our PCR will be a specific locus on chromosome 16 that sometimes contains an Alu sequence. At this site, some individuals have a single Alu sequence inserted. Note that in this case, the presence or absence of the Alu sequence has nothing to do with a disease state and simply serves as a convenient measure of molecular variation. Individuals may have this insert in both homologous chromosomes (+/+), just one homologous chromosome (+/-), or in neither of their homologous chromosomes (-/-). In today’s experiment, we’ll use primers (short regions of single-stranded DNA) that flank the Alu insertion site. These primers will provide the binding site for our DNA polymerase to begin copying our template DNA. If the template DNA does not have the insertion, the amplified fragment will be 641 base pairs (bp) long. If the insertion is present, the amplified fragment will be 941 bp in length.

STUDENT PREPARATION AND GENERAL LAB PROCEDURES FOR THIS LAB.

Before coming to lab, carefully read this lab handout. In addition, review PCR by reading pages 391-392 (Campbell, 7th ed) for Part 1. Before Part 2, read the discussion on Hardy Weinberg Equilibrium, pages 457-458 (Campbell, 7th ed). You should also review the list of terms and concepts outlined under “Objectives”.

OTHER REQUESTS.

Please notify your instructor before beginning this lab if you have a family member who is also in this class.

DETAILED METHODS

Part 1: Extracting Cheek Cell DNA (Day 1)

Our first challenge will be to extract our genomic DNA from our cheek cells. To do this, follow the steps below:

1. Each member of your team should obtain one, screw-cap tube containing 200 ul of the InstaGene matrix and label with their initials or other identifying marks.

2. Obtain a sterile scraper and gently scrape the inside of your cheek ~10 times. Place this scraper in one of the tubes containing the InstaGene matrix. Hold the top of the tube between your thumb and index finger, and vortex briefly to remove the cells from the scraper.

3. Repeat this process with a second, clean scraper. Place the scraper in the same tube of InstaGene matrix and vortex briefly as above.

4. Remove the scraper, cap the tube securely, and shake vigorously.

5. Place your tube in the Styrofoam racks and float the rack in the 56 C water bath for 10 minutes. Approximately halfway through this incubation, remove the tube and shake briefly before returning to the water bath.

6. Remove the tube, shake again, and transfer it to the 100 C water bath for 6 minutes.

7. Remove the tube, shake again, and spin the tube at ~6,000 x g for 5 minutes. (Take care to balance the centrifuge, using equal numbers of tubes on both sides of the rotor.)
Part 2: Amplifying the Target Sequence by PCR (Day 1)

With our DNA successfully isolate from our cheek cells, we’re ready to begin the PCR. We’ll set-up this reaction today and allow it to run overnight. We’ll analyze our samples using gel electrophoresis next week.

1. While your tube is spinning (Part 1, Step 7), label 1 PCR tube (small 200 ul plastic tube) and 1 capless tube. The capless tube will be used to hold your PCR tube in the rack. Note that PCR tubes have thin walls to facilitate heat transfer. This also facilitates the rupture of the tubes, so handle them with care!

2. When your tube has stopped spinning, use a pipette and sterile tip to remove 20 ul of the DNA solution (supernatant) and transfer it to your labeled PCR tube. Take care not to transfer any of the InstaGene matrix!!

3. Locate the tube of PCR Master Mix. (Note that this should be kept on ice!) Remove 20 ul of the yellow PCR Master Mix and add it to the PCR tube. Pipette up and down to gently mix, avoiding air bubbles if possible!

4. Cap the tube tightly (watching those thin walls!) and place it in the thermocycler. The thermocycler will automatically heat and cool our samples, performing 40 cycles of PCR.
AN OVERVIEW OF METHODS

**Day 1**

- Carefully collect cheek cells with a pipet tip
- Add cheek cells to InstaGene matrix in micro test tube
- Vortex or vigorously agitate to extract genomic DNA

- Incubate at 56°C for 10 minutes then repeat vigorous agitation
- Incubate at 100°C for 6 minutes then repeat vigorous agitation
- Centrifuge samples for 5 minutes to pellet matrix

**DNA template preparation**

**Day 2**

- Transfer supernatant with genomic DNA to 200 μl PCR tube

- Add primers to master mix containing:
  - Nucleotides
  - Reaction buffer
  - DNA polymerase

- Add complete master mix to student DNA samples and positive controls
- Place tubes in thermal cycler and amplify target DNA sequence

**DNA amplification**

- The three positive controls supplied in the kit represent the three possible student outcomes.

1. DNA molecular mass ruler
2. Homozygous control (+/+)
3. Heterozygous control (+/-)
4. Homozygous control (-/-)
5. Student sample (+/-)
6. Student sample (+/+)
7. Student sample (+/-)
8. Student sample (-/-)

**Day 3**

Electrophoresis PCR samples on agarose gels at 100 V for 30 minutes. Stain overnight with Bio-Safe DNA stain or ethidium bromide

**Determine student genotypes for Alu insertion and perform Hardy-Weinberg analysis on class results**

**Bioinformatics**
QUESTIONS TO CONSIDER (AND ANSWER)

Day 1:
1. Consider two parents. One parent does not have the insertion in either of version of her chromosome 16. The other parent does have the insert in both versions of his chromosome 16. If you were provided with DNA from their offspring, how many bands would you expect to see if you performed this PCR? Would all offspring have identical banding patterns on the gel? Why or why not?

2. In Part 1, you heated the samples containing your cheek cells. Think carefully about where this DNA was packaged in your cells, and what high temperatures usually do to enzyme structure and function. (Hint: Lysosomes contain enzymes known as DNAses!) What functions do you think these high temperature incubations served in your DNA extractions?

3. In Part 2, you added a “PCR Master Mix” to your DNA to begin the PCR. What ingredients did this PCR master mix have to contain to allow the successful amplification of your DNA?
DAY 2: INTRODUCTION

This week in lab we will continue our study of the Alu insert on Chromosome 16. Recall that last week we isolated our DNA from cheek cells, and used this DNA preparation to amplify a region of Chromosome 16 using PCR. At this site, some individuals have a single Alu sequence inserted. Individuals may have this insert in both homologous chromosomes (+/+), just one homologous chromosome (+/-), or in neither of their homologous chromosomes (-/-). If the template DNA (yours!) does not have the insertion, the amplified fragment will be 641 base pairs (bp) long. If the insertion is present, the amplified fragment will be 941 bp in length.

The size difference in the two PCR products, or DNA fragments, provides a convenient way to distinguish these two molecules. In this week’s lab, we’ll use DNA electrophoresis to distinguish the longer PCR product containing the inserted Alu sequence from the shorter products lacking the insert.

DNA electrophoresis separates DNA fragments according to their size (molecular weight). To do this, DNA fragments are loaded into a slab of agarose gel located in a chamber filled with a conductive buffer. When a current is passed between the ends of the chamber, the negatively charged DNA fragments will be drawn towards the positive pole. The gel acts as a molecular sieve, allowing smaller DNA fragments to move more easily than larger ones. (Thus the rate of travel is inversely proportional to molecular weight.) Over a period of time, smaller DNA fragments will travel farther than larger ones. DNA fragments of the same size stay together, creating discrete bands that can be visualized using a DNA stain.

The pattern of bands we see in the gel will allow us to determine each of our genotypes at this locus. An individual who is homozygous for the insert (+/+ ) should have just a single band visible higher in the gel; this larger (941 bp) molecules will not have traveled as far in the gel. An individual who is heterozygous (+/-) will see two bands on their gel corresponding to the 941 bp product with insert and the shorter, faster, 641 bp product without insert. Similarly, an individual who is homozygous and lacks the insert (-/-) will see just a single band farther down in their gel.

STUDENT PREPARATION AND GENERAL LAB PROCEDURES FOR THIS LAB.

Before coming to lab, carefully read this lab handout and review the handout from last week’s lab. In addition, read the discussion on Hardy Weinberg Equilibrium in your text. You should also review the list of terms and concepts outlined under “Objectives”.

OTHER REQUESTS.

Please notify your instructor before beginning this lab if you have a family member who is also in this class.

DETAILED METHODS

Part 1: Pouring the Gel

1. Obtain an electrophoresis kit from the back counter. Each lab group will need one kit.
2. Remove the tray from the electrophoresis chamber. Use the tape provided to carefully seal each end of the tray.
3. Place the comb in the tray, being sure it fits tightly in the notches on the edge of the tray.
4. Your instructor will provide melted agarose. Handle this solution with care as it is hot! Carefully pour the melted agarose into your tray, filling it up to 2-3 mm past the bottom of the “teeth” on your comb.
5. Allow the gel to cool undisturbed on a level surface for ~15 minutes. The gel is ready when it is slightly opaque and solid to the touch.

Part 2: Loading and Running the Gel

1. Your instructor will direct you to the rack containing your PCR tubes from last week. While your gel is cooling, carefully identify your tube and return to your station. In addition, each lab group will need to obtain one of each of the provided controls (a total of 3 control tubes per group).
2. Carefully tap the PCR tubes to pull any condensation to the bottom of the tubes. Then add 10 ul of the Loading Dye provided in your kit to your sample, mixing gently. This blue loading dye will make our DNA easier to see,
as well as more dense. These properties will be useful to you shortly! Note that the Loading Dye has already been added to the positive controls.

3. When your gel has cooled, remove the tape from the tray and gently pull the comb out of the gel. Place the gel in the electrophoresis chamber with the wells at the negative (black) end. Add ~275 mls of electrophoresis buffer to the chamber until the surface of the gel is covered and both ends of the box are full.

4. When your gel is ready, your instructor will come by and load a molecular size marker in the first lane. This marker will help us determine the size of your PCR products. After this size marker is loaded, load 20 ul of each of your samples in the following order on your gel:

<table>
<thead>
<tr>
<th>Lane</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DNA Size Ladder</td>
</tr>
<tr>
<td>2</td>
<td>Homozygous (+/+) Control</td>
</tr>
<tr>
<td>3</td>
<td>Heterozygous (+/-) Control</td>
</tr>
<tr>
<td>4</td>
<td>Homozygous (-/-) Control</td>
</tr>
<tr>
<td>5</td>
<td>Student 1 Sample</td>
</tr>
<tr>
<td>6</td>
<td>Student 2 Sample</td>
</tr>
<tr>
<td>7</td>
<td>Student 3 Sample</td>
</tr>
<tr>
<td>8</td>
<td>Student 4 Sample</td>
</tr>
</tbody>
</table>

5. Attach the lid to your electrophoresis chamber and plug it into one of the power supplies. Turn on the power supply and set it to 100 V. (Note that this is a relatively high voltage. These chambers are designed to prevent shock, but take care when handling the electrodes!) We will allow the gels to run for ~40 minutes.

Part 3: Staining the Gel
1. When your gel has finished running, carefully unplug it and drain the buffer into the beaker provided.
2. Obtain a staining dish, label it, and transfer your gel to the dish. (Note that these gels are both slippery and fragile!)
3. Pour 100X Fast Blast DNA stain into the staining dish until it covers the surface of your gel. This stain contains a positively charged blue dye with a strong affinity for DNA. Allow the gel to stain for ~3 minutes.
4. Pour the stain off into the beaker provided. We will now wash any excess dye out of the gel.
5. Fill your staining dish with warm water. Swirl gently for ~10 seconds and pour the water down the sink.
6. Transfer your gel to the larger container provided and fill it with warm water. Continue to wash the gel by allowing it to sit in this water for 5 minutes, with occasional mixing. While your gel is washing, rinse out your electrophoresis chambers, dry them carefully and return them to your kit!
7. Replace the water in this container and repeat step 6. (This wash may be repeated a third time if the background staining is still high.)
8. When washing is complete, return your gel to the staining dish. The bands may appear fuzzy at first, but should sharper over the next 5-10 minutes as the Fast Blast dye molecules migrate and bind more tightly to the DNA.

Part 4: Analyzing the Gel
1. When you are satisfied with the appearance of your gel, your instructor will help you take a picture of the image. Work from this picture to analyze your results and record them in the class data table.

ANALYZING THE RESULTS
Compare your sample lanes with the control lanes, using the DNA size marker as a reference. By comparing your DNA migration pattern to the controls, you should be able to determine whether you are homozygous +/-, heterozygous +/-, or homozygous -/-.

Work from this picture to analyze your results and record them in the class data table. When this class table is complete, copy the result to the table below:

My genotype: ___________________________
Now for the fun part! Because these Alu repeats have become integrated into the general population at random, the 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 some others the individuals may all be homozygous -/-, while in a “melting-pot” population the three genotypes 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 our class.

Recall that the Hardy-Weinberg equation (below) describes the frequencies of alleles in the gene pool of an entire population. In this case the entire population is our class.

\[ p^2 + 2pq + q^2 = 1 \]

By determining frequencies of the Alu genotype within your class population, you can calculate the allelic frequencies. Additionally, the genotypic frequencies of our class population can be compared to published results of larger population sizes.

Use the class genotype frequency (Table 2) to calculate the allelic frequencies for the class. Remember that a class of 32 students will have a total of 64 alleles. Record the calculated frequencies in the table below.

### Table 3: Class Allelic Frequency

<table>
<thead>
<tr>
<th>Category</th>
<th>Number</th>
<th>Class Allelic Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (+) alleles</td>
<td>#/Total = = p</td>
<td></td>
</tr>
<tr>
<td>Total (-) alleles</td>
<td>#/Total = = q</td>
<td></td>
</tr>
<tr>
<td>Total alleles =</td>
<td>( p + q = 1 )</td>
<td></td>
</tr>
</tbody>
</table>

Once you’ve calculated the allelic frequencies, we can then compare ourselves to the results of a USA-wide random population study. In a random sampling, researchers have observed the following genotypic frequencies in the US population:

### Table 4: Sampled USA Genotypic Frequencies

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of each genotype</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homozygous (+/+)</td>
<td>2,422</td>
<td>0.2422</td>
</tr>
<tr>
<td>Heterozygous (+/-)</td>
<td>5,528</td>
<td>0.5528</td>
</tr>
<tr>
<td>Homozygous (-/-)</td>
<td>2,050</td>
<td>0.2050</td>
</tr>
<tr>
<td>Number of Samples</td>
<td>Total = 10,000</td>
<td>Total = 1.0</td>
</tr>
</tbody>
</table>

You can use this data to then determine the allelic frequencies for the general US population just as you did for our own data in Table 3. Calculate the allelic frequencies for the US population and record them in Table 5 below.

### Table 5: US Population Allelic Frequency

<table>
<thead>
<tr>
<th>Category</th>
<th>Number</th>
<th>US Allelic Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (+) alleles</td>
<td>#/Total = = p</td>
<td></td>
</tr>
<tr>
<td>Total (-) alleles</td>
<td>#/Total = = q</td>
<td></td>
</tr>
<tr>
<td>Total alleles =</td>
<td>( p + q = 1 )</td>
<td></td>
</tr>
</tbody>
</table>
QUESTIONS TO CONSIDER (AND ANSWER ON A SEPARATE PAGE)

4. Examine the gel of your PCR products. Based on its appearance, do you think your reaction (PCR) worked as expected? Why or why not?

5. Compare our class data for allelic frequencies with that of the random US population. Do they match? Do you expect them to match? What reasons can you think of to explain the differences or similarities?

6. Use the calculated allelic frequencies of our class to determine the expected genotypic frequencies if our population is in Hardy Weinberg equilibrium. (Show your work!) How do these expected genotypes compare to those we actually observed? If there are differences, which conditions of the Hardy-Weinberg equilibrium have we most likely violated?

To Turn In:

Answers to the questions (from both days!)
Complete tables 2, 3, and 5