

LABORATORY MODULE 4 FOR BIOLOGY 273

Population Genetics: Analysis of Alu Insertion Polymorphism

(Portions of this module are produced and copyrighted by the Dolan DNA Learning Center for noncommercial, educational use only)

Orientation

A. Purpose: The purpose of this module is to acquire a basic understanding of population genetics. At the conclusion of the module, students should understand: 1) the concept of transposable elements; 2) the concepts of polymorphisms and genetic variation; 3) the concepts of Hardy-Weinberg Equilibrium and genetic drift. Specifically, the students will determine the allelic frequency of an *Alu* insertion in a specific locus in the student population.

B. Reflection on Prior Knowledge: Discuss the following questions with your lab group then write a brief answer to each of the following questions in your lab notebook:

- *What are "transposable DNA elements"?*
- *What factors can influence the frequency with which an allele exists within a population?*
- *Do all changes to the DNA sequence of a gene affect the function of the gene?*

C. Timeline:

	Week 1	Week 2	Week 3
Activity	<u>Session 1</u> <ul style="list-style-type: none">• Introduction• Experiment Day I <u>Session 2</u> <ul style="list-style-type: none">• Experiment Day II• Concepts: activity and work on problems with Hardy-Weinberg Equilibrium (5pts)	<u>Session 3</u> <ul style="list-style-type: none">• <u>Lab practical (5pts)</u>• Experiment Day III• Work on Exercises I-III <u>Session 4</u> <ul style="list-style-type: none">• Explore Cold Spring Harbor Lab website for <i>Alu</i> comparison group in preparation for paper.	<u>Session 5-6:</u> Organize and gather data for paper on affects of genetic drift and selection on world populations

D. Connections: Population genetics is one aspect of the larger field of population biology. Population biology is the basis of the Biology Department's Biodiversity and Ecology and Conservation Biology emphasis. There are a number of researchers at CSUF who study organism populations including Drs., Eernisse, Banack, Jones, Murray and Sandquist. (See <http://biology.fullerton.edu> for details of their research). This module utilizes one technique with which these researchers can determine the migration patterns or ancestry of a population of animals using one genetic marker, the presence (or absence) of an *Alu* element insertion on human chromosome 16. Outside of academia, the practice of "DNA fingerprinting" crime suspects for genetic markers, such as *Alu* insertions, has also been proven useful in determining the guilt or innocence of crime suspects.

E. Additional Resources

- Any Genetic text that covers population genetics topics of allele frequency, Hardy-Weinberg Equilibrium and genetic drift and natural selection is also appropriate
- <http://www.dnafb.org/dnafb/32/concept/index.html> ('animations' link). *Details about Barbara McClintock's work with transposable elements.*
- <http://www.woodrow.org/teachers/bi/1994/hwintro.html>. *A good source for activities with Hardy-Weinberg Equilibrium.*

Introduction (Excerpts from Cold Spring Harbor Laboratory module)

Population Genetics

By all accounts, the road to the evolution of our species began in Africa. Fossils of modern humans, dating to 40,000-100,000 years ago, have been found throughout the "Old World" – Africa, Europe, and Asia – and in Australia. By modern humans, we mean members of our own species, *Homo sapiens*, who shared with us important anatomical features. These humans subsequently spread to Micronesia, Polynesia, and North and South America.

The means of tracing the migration of populations, both human and non-human, has been made simpler by the advent of molecular biology. Using molecular biology techniques, one can readily get an idea of the ancestry of a sub-population. This can be done by examining a defined set of chromosome positions (called "loci", plural of "locus"). Subtle sequence changes in the DNA of multiple loci can give a "genetic fingerprint" of an individual. This can allow an individual to be either grouped with others with a somewhat similar genetic fingerprint or singled out by having some unique differences.

On the other hand, one can focus on one gene and study the frequency with which it occurs in a population -- the basis of Hardy-Weinberg Law (HWL). HWL began as a mathematical model devised by Godfrey Hardy and Wilhelm Weinberg in 1908 to explain the effects of mating on allele and genotype frequencies in a population. Basically HWL assumes that if a population is large, randomly mating, and has negligible mutation, migration or natural selection then: 1) allele frequencies in the population will remain constant; 2) genotypic frequencies will remain constant after one generation of random mating. That is, the population that meets the criteria of the assumptions of HWL will have allelic and genotypic frequencies in *equilibrium* within the population. The two alleles of this module is presence ("+") or absence ("-") of an *Alu* element insertion at one location ("PV92") on human chromosome 16.

The *Alu* element

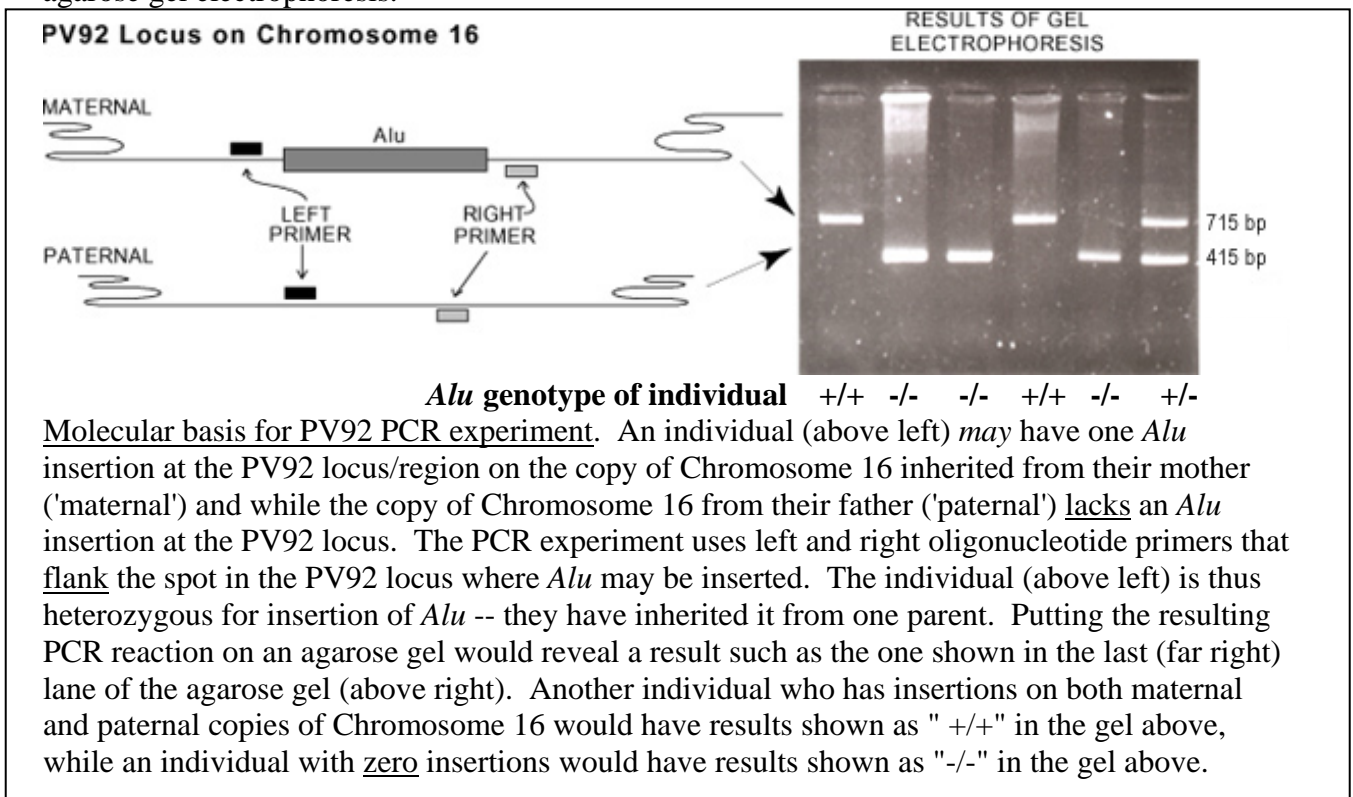
The *Alu* segment is an example of a so-called "jumping gene" – a transposable DNA sequence that "reproduces" by copying itself and inserting into new chromosome locations. There are thousands of copies of *Alu* insertions in the human genome. The actual number of copies can vary between individuals. All *Alu* elements are approximately 300 bp in length and derive their name from a single recognition site for the restriction enzyme *AluI* located near the middle of the *Alu* element. *Alu* is a "defective" transposable element, in that it lacks the enzyme functions to produce a DNA copy of itself and to integrate into a new chromosome position. Nobel laureate Barbara McClintock hypothesized that transposable elements provide a mechanism to rapidly reorganize the genome in response to environmental stress.

It is important to note that most *Alu* mutations are "fixed," meaning that both of the paired chromosomes have an insertion at the same locus (position). However, a number of human-specific *Alus* are *dimorphic* -- that is, an insertion may be present or absent on each of the

paired chromosomes of different people. These dimorphic *Alu* elements inserted within the last million years, during the evolution and dispersion of modern humans. These dimorphisms show differences in allele and genotype frequencies between modern populations and are tools for reconstructing human prehistory.

Once an *Alu* inserts at a chromosome locus, it can copy itself for transposition, but there is no evidence that it is ever excised or lost from a chromosome locus. So, each *Alu* insertion is stable through evolutionary time. Each is the "fossil" of a unique transposition event that occurred only once in primate evolution. Like genes, dimorphic loci (that may or may not have an *Alu* insertion) are inherited in a Mendelian fashion from parents to children.

The experimental portion of this module examines PV92, a human-specific *Alu* insertion site on chromosome 16. The PV92 locus will have only two alleles: the presence (+) or absence (-) of the *Alu* transposable element on each of the paired chromosomes. This results in three PV92 genotypes (+/+, +/-, or -/-). The + and - alleles can be determined using PCR followed by agarose gel electrophoresis.



Laboratory Exercises

Session 1

Experiment Day I: DNA Isolation by Saline Mouthwash

Reagents	Equipment & Supplies	Shared Items
<ul style="list-style-type: none">• Saline solution (0.9% NaCl), 10 ml/student in sterile 15ml tubes.• Paper cups• one 1.5 ml microcentrifuge tube per student• 10% Chelex®, pH 11, 100 µl per student.	<ul style="list-style-type: none">• 1.5 ml microcentrifuge tubes• 15ml conical tube• P1000 micropipettor• P200 micropipettor	<ul style="list-style-type: none">• Microcentrifuge• Boiling water bath

Procedure

1. Pour 10 ml of the saline solution (0.9% NaCl) into mouth and vigorously swish for 30 seconds.
2. Expel saline solution into a paper cup.
3. Swirl to mix cells in the cup and transfer 1 ml (1000 µl) of the liquid to 1.5 ml tube using P1000 micropipettor.
4. Place your sample tube, together with other student samples, in a balanced configuration in a microcentrifuge and spin at full speed (about 13,000rpm) for 1 minute.
5. Carefully pour off supernatant the marked waste container. DO NOT dump into the sink. **Your wash is considered to be biological waste and must be disposed of properly.** Be careful not to disturb the cell pellet at the bottom of the test tube. Alternatively, you may use a P1000 micropipettor and a new tip to do this. A small amount of saline will remain in the tube.
6. Resuspend cells in remaining saline by pipetting in and out with a P200 pipettor. (If needed, 30 µl of saline solution may be added to facilitate resuspension.)
7. Withdraw 30 µl of cell suspension, and add to tube containing 100 µl of Chelex. Shake well to mix.
8. Boil cell sample for 10 minutes. Use boiling water bath. Then, cool tube briefly on ice for about 1 minute. This cuts down on any vapor pressure inside the tube that may have accumulated during boiling.
9. Shake tube by flicking bottom of tube with your fingers. Place in a balanced configuration in a microcentrifuge, and spin for 1 minute.
10. Transfer 30 µl of supernatant (containing the DNA) to clean 1.5 ml tube. Avoid cell debris and Chelex beads. This sample will be used for setting up one or more PCR reactions.

11. Store your sample in the freezer until ready to begin Part II in session 2.

SESSION 2

Experiment Day II: DNA Amplification by PCR

Reagents	Equipment & Supplies	Shared Items
<ul style="list-style-type: none">• Primer mix, 22.5 μl• Human DNA, 2.5 μl• Ready-to-Go Bead (in reaction tube)	<ul style="list-style-type: none">• P20 and P200 micropipettors and tips	<ul style="list-style-type: none">• Thermal Cycler

1. Use a micropipette with a fresh tip to add 22.5 μ l of the primer mix to a PCR tube containing a Ready-To-Go PCR Bead. The primer mix contains water and oligonucleotide primers that bind to sequences flanking the PV92 insertion site. The Ready-to-Go bead contains PCR buffer, dNTPs and taq DNA polymerase. Thus, all constituents are here for a PCR reaction, except for template DNA.

2. Use fresh tip to add 2.5 μ l of human DNA (from Part I) to reaction tube, mixing with your pipette tip. Pool reagents by sharply tapping tube bottom on lab bench.

3. Label the cap of your tube with a number, as assigned by your teacher. In this way, your results will be anonymous.

4. Store all samples on ice until ready to amplify according to the following profile. Program thermal cycler for 30 cycles according to the following cycle profiles. Each program may be linked to a 4°C to hold samples after completing the cycle profile, but amplified DNA also holds well at room temperature.

Denaturing time and temperature 30 sec, 94°C

Annealing time and temperature 30 sec, 58°C

Extending time and temperature 30 sec, 72°C

Concepts in Population Genetics: Hardy-Weinberg Equilibrium

(reference (with permission): Judith Stanhope, Woodrow Wilson Biology Institute)

In this lab module, you are determining the allelic frequency of an *Alu* insertion in a specific locus in the student population. This touches on a key concept in population biology -- the importance of measuring the frequency of alleles within a population. Why is this so important? Thomas Merten (*The American Biology Teacher*, vol 54, no. 2. pp. 103-107 1992) said that the importance of population biology is illustrated by these questions:

"How can type O blood be the most common blood type if it's a recessive trait?"

"If Huntington's disease is a dominant trait, shouldn't 3/4 of the population have Huntington's while 1/4 have a normal phenotype?"

"Shouldn't recessive traits be gradually "swamped out" so they disappear from the population?"

These questions reflect the common misconception that the dominant allele of a trait will always have the highest frequency in a population and the recessive allele will always have the lowest frequency. Gene frequencies can be high or low no matter how the allele is expressed, and can change, depending on the conditions that exist. It is the changes in gene frequencies over time that result in evolution.

It is at the population level that evolution is evident. A population can be thought of as a group of individuals of the same species in a given area who can interbreed. Because of this interbreeding, they share a common set of genes known as a gene pool. This gene pool contains *all* the alleles for *all* the traits of *all* the population. For evolution to occur in real populations, some of the gene frequencies must change with time.

An important way of discovering why real populations change with time is to construct a model of a population that *does not* change. This is what Hardy and Weinberg did. Their principle describes a situation in which there is no change in the gene pool (the frequencies of alleles), and thus no evolution.

Let's focus on one gene in the gene pool that has alleles A and a . Hardy and Weinberg assigned the letter p to be the frequency of the dominant allele in the pool and q to be the frequency of the recessive allele in the pool. Since the sum of the alleles must equal 100%,

$$p + q = 1.00$$

The allele frequency , for one allele, can be thought of as:

$$\begin{aligned} \text{allele frequency for one allele} &= \frac{\text{the number occurrences of that allele in the gene pool}}{\text{the total number of alleles in the gene pool}} \\ &= \frac{2(\text{number of homozygotes for that allele}) + (\text{the number of heterozygotes})}{\text{the total number of alleles in the gene pool}} \end{aligned}$$

Ex) In a gene pool of 30 alleles (A or a): the number of A alleles =12 and the number of a alleles=18.

$$\text{frequency}(A) = 12/30 = 0.4 = 40\%$$

$$\text{frequency}(a) = 18/30 = 0.6 = 60\%$$

$$\text{Notice that, } p + q = 0.4 + 0.6 = 1.00$$

Moreover, Hardy and Weinberg noticed that all the combinations of alleles (AA or Aa or aA or aa) could be represented in terms of allele frequencies:

$$p^2 + pq + qp + q^2, \text{ (rewritten as } p^2 + 2pq + q^2 \text{)}$$

Ex) (Using A and a frequencies in previous example)

$$p^2 + 2pq + q^2 = AA + 2Aa + aa = (0.4)^2 + 2(0.4)(0.6) + (0.6)^2$$

$$= .16 + .48 + .36 = 1.00 \text{ (or 100\%)}$$

Notice that this can tell you the percentage of a genotype in the population. That is, there are 16% AA, 48% Aa and 36% aa.

To keep the allele frequencies constant, Hardy and Weinberg noted that NO changes to the Hardy-Weinberg gene pool occur (and allele frequencies remain unchanged generation after generation) if the following conditions are met:

- 1) Large population. The population must be large to minimize random sampling errors.
- 2) Random mating. There is no mating preference such as an AA male must not prefer an aa female.
- 3) No mutations occur. Alleles must not change
- 4) No migration occurs. No exchange of genes between different populations
- 5) No natural selection occurs. Natural selection must not favor any individual

For the remainder of session 2, you will participate in an activity (instructions given by the instructor) that illustrates the principle of Hardy-Weinberg Equilibrium. You will also work on an in-class problem set that is worth 5pts toward your total lab grade.

Session 3

Experiment Day III: DNA Analysis by Gel Electrophoresis

Reagents	Equipment & Supplies	Shared Items
<ul style="list-style-type: none"> • 2% agarose gel stained with ethidium bromide • 15 ml test tube, polypropylene • 1X electrophoresis buffer • 1kb ladder marker 	<ul style="list-style-type: none"> • 1-20 µl micropipettor and tips 	<ul style="list-style-type: none"> • Electrophoresis chamber • Electrophoresis power supply

Procedure

1. Use a micropipette with a fresh tip to add 5ul of loading dye to your PCR sample. Next add as much PCR sample/loading dye mixture as will fit into your assigned agarose gel well. Expel any air from the tip before loading, and be careful not to push the tip of the pipet through the bottom of the sample well.

2. Load 5 µl of the "1kb ladder" size markers into one lane of gel.

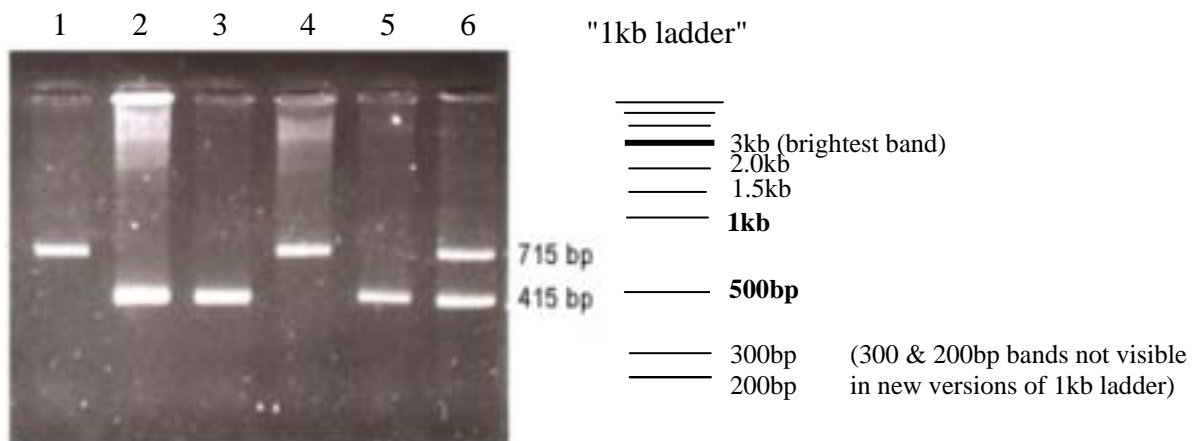
3. Electrophorese at 110 volts for 45-60 minutes. Adequate separation will have occurred when the dye front has moved at least 50 mm from the wells.

EXERCISES I-III (answer questions/problems in **bold and underlined text** on a separate sheet of paper that you must turn in to your instructor at the end of session 3).

Exercise I. Interpreting Your Gel

1. Observe the photograph of the stained gel containing your sample and those from other students. Orient the photograph with the sample wells at the top. Interpret the band(s) in each lane of the gel:

a. Scan across the photograph to get an impression of what you see in each lane. You should notice that virtually all student lanes contain one to two prominent bands. See sample below.



b. Now locate the lane containing the DNA markers. Working from the well, locate the bands corresponding to each DNA fragment.

c. One band visible. Compare its migration to that of the 1kb and 500bp bands in the 1kb marker lane. If the PCR product migrates just below the 1kb band (lane 4), then the person is homozygous for the PV92 *Alu* insertion (+/+). Note that the visible band is actually two bands of the same size. If the PCR product migrates below the 500bp (lane 2), then the person is homozygous for the absence of the PV92 *Alu* insertion (-/-). Again, the visible band is actually two bands of the same size.

d. Two bands visible. Compare its migration to that of the 1kb and 500bp bands in the 1kb ladder lane. Confirm that one PCR product corresponds to a size of about 715-bp and that the other PCR product corresponds to a size of about 415-bp (lanes 1 and 3). The person is heterozygous for the PV92 *Alu* insertion (+/-).

e. It is common to see an additional band lower on the gel. This diffuse (fuzzy) band is theorized to be "primer dimer," an artifact of the PCR reaction that results from the primers overlapping one another and amplifying themselves. Primer dimer is approximately 50 bp. (The presence of primer dimer confirms, in the absence of other bands, that the reaction contained all components necessary for amplification, but that there was insufficient template to amplify the PV92 locus.)

f. Additional faint bands at other positions occur when the primers bind to chromosomal loci other than PV92 and give rise to "nonspecific" amplification products. **Why might this occur?**

2. Determine the genotype distribution for the class by counting the number of students of each genotype (+/+, +/-, and -/-).

3. What can you say about the genetic origin of someone in class who has an *Alu* insertion (a '+' allele)?

4. An allele frequency is a ratio comparing the number of copies of a particular allele to the total number of alleles present. Imagine a class of 100 students that list their genotype distribution as follows:

genotype	+/+	+/-	-/-
number in class	20	50	30

Since humans are diploid, the total number of alleles in the class is $2 \times 100 = 200$.

The allele frequency for PV92+ is:

$$2 \times 20 \text{ (homozygotes)} + 50 \text{ (heterozygotes)} / 200 = 90 / 200 = 0.45$$

Likewise, the allele frequency for PV92- is:

$$2 \times 30 \text{ (homozygous)} + 50 \text{ (heterozygotes)} / 200 = 110 / 200 = 0.55$$

Using the genotype distribution from your class, calculate the frequencies of the + and - alleles of PV92.

5. Recall from Hardy-Weinberg equilibrium that the distribution of genotypes for an allele are described by the equation: $p^2 + 2pq + q^2 = 1.00$ where p and q represent the allele frequencies; p^2 and q^2 are the homozygote frequencies; and $2pq$ is the heterozygote frequency. **Use the allele frequencies calculated for your class in Step 4 to determine the expected genotype frequencies. How do they compare with the actual genotype frequencies? How can you account for differences?**

Exercise II. Is Your Class In Hardy-Weinberg Equilibrium?

Differences between observed and expected genotypes may be due to chance or may indicate that the sample population is not in Hardy-Weinberg equilibrium. To help decide which of these explanations is more likely, one uses a statistic test called the Chi-Square. **Use the Student**

Allele Server at the DNA Learning Center WWW site to test Hardy-Weinberg equilibrium for your class.

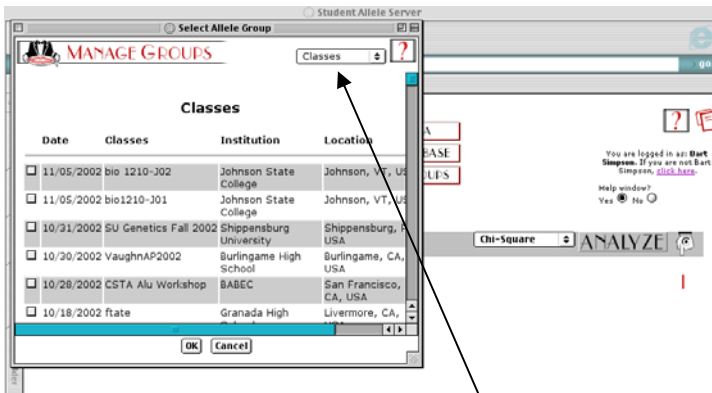
1. Go to the login page (http://www.bioservers.org/sad/login.pl?no_auto=1)



2. In the upper right corner, log in by typing in: 'Bio273' under *username* and 'csuf' under *password* and click on the 'log in' link. You should get the window below



3. If your TA has not pre-loaded the allele server with your group (shown as 'bio273_[your TA]'), then Click on 'MANAGE GROUPS.' When the window opens, WAIT UNTIL THE "CLASSES" LIST COMPLETELY FINISHES LOADING which may take 30 seconds.



Use the pull-down menu to select '**Your Groups.**' Then click on the box to the left of your group name from the different sections listed (again shown as 'bio273_[your TA]').

4. CLICK 'OK'. About 10 sec. later, your lab section will appear upon the Allele Server desktop
5. To see if your group is in Hardy-Weinberg equilibrium, choose 'Chi-Square' from the pull-down menu on the RIGHT SIDE of the screen. Click on the round button to the right side of the screen (where the 'hand' is pointing) to select your group for analysis. Then click 'ANALYZE.'
6. The Allele Server will perform a Chi-Square analysis using your observed genotype frequencies and the genotype frequencies predicted from a population in Hardy-Weinberg equilibrium. The predicted genotype frequencies are calculated using the same approach you used in Session 3, 'Experiment Day I'.

The analysis page shows pie charts for the observed and expected genotype frequencies. **Do these pie charts look substantially different?** The Chi-Square test provides a statistical measure for the difference between the two sets of frequencies. In general, the higher the Chi-Square value, the greater the difference between the observed versus the expected frequencies. *Clicking on the 'verbose' tab gives you detailed information about the pie charts.*

The Chi-Square test for Hardy-Weinberg equilibrium assumes the "null hypothesis" - that is, the observed genotype frequencies are not significantly different from those predicted for a population in equilibrium. As with any theoretical value, the genotype frequencies predicted by the equilibrium equation almost always differ from the frequencies observed in a real population. The problem is to discern when the observed versus expected values differ due to chance and when they are truly different. Recall that the actual calculation of Chi-Square for a given phenotype is:

$$(\text{observed frequency} - \text{expected frequency})^2 / \text{expected frequency}$$

The Chi-Square values for each phenotype are then added together and degrees of freedom are taken into account.

A probability value, or p-value, is used to evaluate the significance of a Chi-Square. Scientists give a wide margin for differences that may occur by chance by setting the cutoff for significance at p-value of 0.05 (5%) or less. This means that one may expect a Chi-Square of this value to occur by chance in 5% of genotype comparisons. Conversely, there is a 95% probability that the differences between observed versus expected genotype frequencies are "real." Social scientists expand the probability window by saying that p-values between 0.5 and 0.10 "approach significance."

For example, a p-value of 0.34 means that there is a 34% probability that the genotype *differences* are due to chance and 66% chance that they are *not* due to chance. This p-value is not significant, the null hypothesis is upheld, and we say that the population is in Hardy-Weinberg equilibrium. A p-value of 0.02 means that there is a 2% probability that the genotype differences

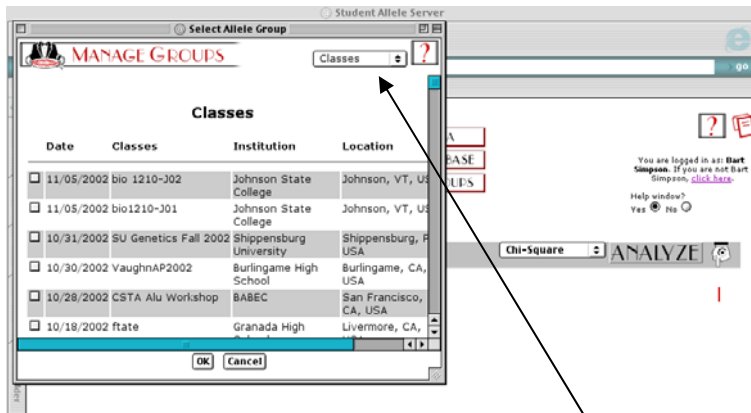
are due to chance and 98% chance that they are not due to chance. This p-value is significant, the null hypothesis is rejected, and we say that the population is not in Hardy-Weinberg equilibrium.

What is the p-value for your Chi-Square? Is it less than .05? If so, can you suggest any factors that might account for why your observed population is not in Hardy-Weinberg equilibrium?

Exercise III. How Does Your Class Compare to Other Groups?

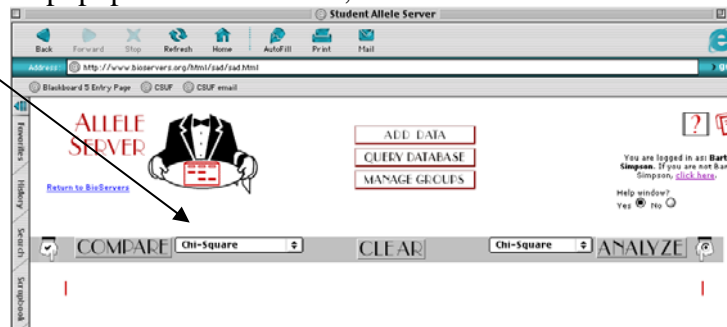
1. *Alu* insertion frequencies can also be used to study relationships between human populations. The Chi-Square statistic can be used to determine if *Alu* genotype frequencies differ significantly between any two populations. **In this section, you will use the Student Allele Server to compare your observed *Alu* genotype frequencies to reference data sets collected from populations around the world.**

a. If the Allele Server is not already on your screen, follow Exercise II, steps 1-4 again.



b. Click on 'Manage Groups' again. Choose "Reference" from the popup menu. Click the checkboxes next to four or five reference groups drawn from around the world. Click OK. The Manage Groups window closes, and the populations you selected are loaded into the Allele Server workspace.

c. On the Sequence Server Workspace, click the checkbox next to your workshop, and then click the checkbox next to any one of the reference sets you chose. Choose "Chi-Square" from the popup menu on the left, then click 'COMPARE.'

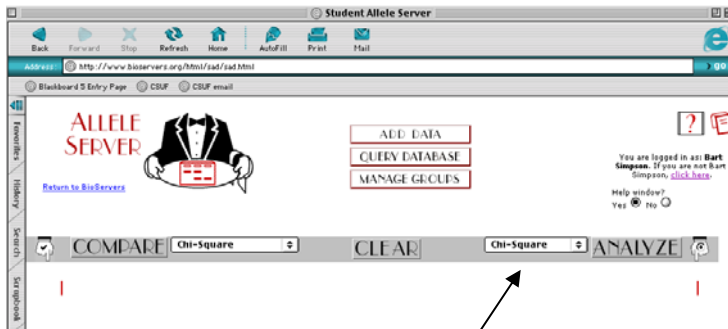


d. On the Chi-Square page, compare the genotype frequency pie charts for the two populations. Do the populations look similar or different? Does the Chi-Square statistic and associated p-value support your visual impression?

e. Click DONE to close the Results window. Choose another reference data set from the Workspace, and run the Chi-Square test again. Record your results in a notepad.

f. To complete your analysis, uncheck your workshop, then begin checking other reference sets. Repeat these pairwise comparisons until all permutations have been performed.

g. Based on the results you recorded, how useful is the PV-92 *Alu* polymorphism in distinguishing populations from each other? Why do you think that the PV-92 allele frequencies differ significantly between some populations, while not between others? Do you think you could use PV-92 data from a large number of populations to answer the questions of where humans originated and the paths by which they spread throughout the world? What other population data sets or types of information might you need to accurately answer this question?



2. Use the Heterozygosity Feature of the Analyze Function to determine the + allele frequency in a number of populations representing differing parts of the world. Select 'Heterozygosity' from the pull down menu on the RIGHT. Click the round checkbox underneath to select a group for analysis. (You can only analyze one group at a time). Detach the picture of the world map in this module (also found at <http://www.dnalc.org/map.html>) and plot the + allele frequencies on it.

a. Do you notice any pattern in the allele frequencies?

b. Suggest a hypothesis about the origin and dispersal of the *Alu* allele that accounts for your observation.

c. After comparing *Alu* allele frequencies among different world populations, do you think this procedure of PCR amplification of *Alu* insertion sequences could be used forensically to link a suspect with a crime? Could it be used if the suspects were from different parts of the world?

Sessions 4-6

Preparation for presentation on population genetics (50pts).

*Learning outcome: to discover how not fulfilling the assumptions of Hardy Weinberg Law can throw a population that was once at Hardy-Weinberg equilibrium for allelic and genotypic frequencies into **changed** allelic and genotypic frequencies for the population. In short, to simulate evolutionary forces that can change the genetic makeup of a population.*

Earlier in this module (session 2) it was stated that one can begin to study why populations change over time by studying a population in Hardy-Weinberg equilibrium that does *not* change over time. We will use the Cold Spring Harbor site to now study a population that **will** change over time because it does not meet all of the assumptions of a population in Hardy-Weinberg equilibrium.

In Lab sessions 4-5, your group will take the results of the class (+) allele frequency (the frequency of having an *Alu* insertion) and compare this (+) allele frequency to that of a number of comparison groups of *your choosing*. (You may wish to pick the comparison group from the world map on which you noted the '+' allele frequency of a number of populations. You will then use a Hardy-Weinberg simulator that illustrates effects of genetic drift (basically, subtle and temporary changes to the allele frequency of a *small* population over time) and evolution (more permanent changes, i.e. over many generations, to the allele frequency of a population over time).

The key to getting interesting results is to set the population of each group to a small number (50 or less people). Genetic drift, again, consists of small changes in the allele frequency over time. With a large population already at Hardy-Weinberg equilibrium, these changes don't have much of a permanent affect. With small populations (such as a class of 24 students), small changes in allele frequency can have drastic effects over time. These effects are further accentuated if an allele decreases the chance for survival of an individual (also known as selection).

The Hardy-Weinberg simulator allows you to study Hardy-Weinberg population equilibrium and how it is affected by changes in genotype survival rates, initial allele percentages, population size, and the number of generations. At each generation, two parents are chosen at random from a population and a child's genotype is generated from its parents' genotypes, using an approach similar to performing a Punnett Square analysis. The probability that the child survives to the age of reproduction is determined by the survival rate for the child's genotype. If the child does survive, then that child is added to the gene pool for the next generation. This process is repeated, starting with choosing parent pairs, until enough children are generated so that the number surviving is the same as the number of parents.

Through the computer simulation, YOU decide such things as the genotype survival rates, starting allele frequency, and starting population size in one or more of your populations in a scenario that you determine. The simulator can even show what can happen if two

populations are mixed ('linked') -- although this option puts a load on the server and is NOT recommended.

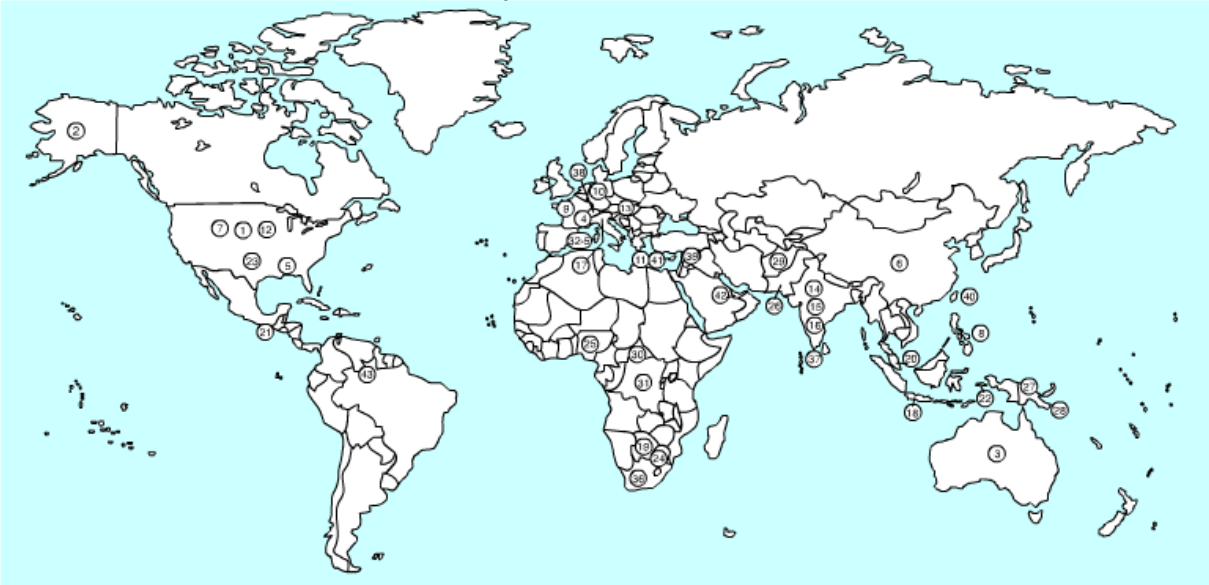
Example scenario) *The Alu PV92 insertion is found to result in a subtle immunity against a newly characterized disease. The World Health Organization wants you to give a report on the possibility of survival of three small populations (from different parts of the world) over the course of the foreseeable future (i.e. 100 generations).*

Your TA will give you a mini-manual on how to run the simulator

Your Power Point presentation must:

- Analyze at least 3 different populations (one may be the 273/320L class) using the '+' allele frequencies that are either on the Allele Server (for the 'reference groups' of world populations) or that were determined by your group (for the class allele frequency). Note, again, that you wrote down the '+' allele frequencies of these populations in Exercise III.
- Provide some background on these populations: location of the population in the world, pie chart or graphic of the frequency of *Alu* insertions in that population, reason they relate to your scenario and possibility for outside contact or 'mixing' of the population.
- Provide some scenario for selecting the survival rate of the (+) or (-) allele over 100 generations.
- Talk about the parameters you put into the simulator program and why
- Show your results in the simulator (using at least 3 trials). Discuss why there might be differences in results between the trials.
- Formulate a plan to address the scenario topic based on your simulator results (e.g. *quarantine population X to their part of the world because they have are likely to spread the disease*).

World map (for plotting '+' allele frequencies).



- | | | | | | |
|---------------------|-----------------------|------------------------|--------------------------------------|-----------------------------|--------------------------|
| 1. African American | 9. French | 17. Italian, Northern | 25. Nigerian | 32. Sardinian (Aritzo) | 40. Taiwanese |
| 2. Alaskan Natives | 10. German | 18. Java | 26. Pakistani, Balzer Data | 33. Sardinian (Marrubiu) | 41. Turkish, Cyprus |
| 3. Australian | 11. Greek, Cyprus | 19. 'Kung' ('Bushmen') | 27. Papua New Guinea | 34. Sardinian (Ololai) | 42. United Arab Emirates |
| 4. Breton | 12. Hispanic American | 20. Malay | 28. Papua New Guinea, Costal | 35. Sardinian (San Teodoro) | 43. Yanomamo |
| 5. Cajun | 13. Hungarian | 21. Maya | 29. Pushtoon (Afghan) | 36. Sotho | |
| 6. Chinese | 14. India Christians | 22. Moluccan | 30. Pygmy (Central African Republic) | 37. South India | |
| 7. Euro-American | 15. India Hindu | 23. Mvekoke | 31. Pygmy (Zaire) | 38. Swiss | |
| 8. Filipino | 16. Indian Muslim | 24. Nguni | | 39. Syrian | |

GLOSSARY

Use these URLs to look up definitions.

<http://www.genome.gov/glossary.cfm>

At this site sponsored by the NIH, you will be able to look up a multitude of classical and molecular genetic terms. There is audio real player, which can be downloaded for free, which will allow you to listen to explanations of the terms by world-renowned scientists. Also there is free adobe acrobat, which will allow you to view numerous pictures.

<http://helios.bto.ed.ac.uk/bto/glossary/>

A very comprehensive genetics glossary: from the classical to molecular!