Chapter 5

Demystifying Hardy-Weinberg: Using Cellulose Acetate Electrophoresis of the *Lap* Locus to Study Population Genetics in White Campion (*Silene latifolia*)

Patricia A. Peroni & David E. McCauley

Patricia A. Peroni
Department of Biology, Davidson College
P.O. Box 1719
Davidson, NC 28036-1719
Phone: (704) 892-2890
email: paperoni@davidson.edu

David E. McCauley
Department of Biology, Vanderbilt University
Nashville, TX 37235

Patricia A. Peroni received a B.A. in Biology from the SUNY College at Plattsburgh, an M.S. in Biology from Bucknell University, and a Ph.D. in Botany from Duke University. She is Associate Professor of Biology at Davidson College where she teaches courses in introductory biology, ecology, biostatistics, and environmental issues. She studies the effects of dormant seed banks on the dynamics of white campion (*Silene latifolia*) metapopulations.

David E. McCauley received a bachelor's degree in Zoology from the University of Maryland at College Park and a Ph.D from SUNY Stony Brook. He is Professor of Biology at Vanderbilt University where he teaches biostatistics, evolution, and population biology. His research investigates the ecological genetics of insects and plants with special reference to the effects of population structure on micro-evolutionary processes.

Address correspondence to P.A. Peroni

Reprinted From: Peroni, P. A. and D. E. McCauley. 1999. Demystifying Hardy-Weinberg: Using cellulose acetate electrophoresis of the Lap locus to study population genetics in white campion (*Silene tatifolia*). Pages 101-122, *in* Tested studies for laboratory teaching, Volume 20 (S. J. Karcher, Editor). Proceedings of the 20th Workshop/Conference of the Association for Biology Laboratory Education (ABLE), 399 pages.

- Copyright policy: http://www.zoo.utoronto.ca/able/volumes/copyright.htm

Although the laboratory exercises in ABLE proceedings volumes have been tested and due consideration has been given to safety, individuals performing these exercises must assume all responsibility for risk. The Association for Biology Laboratory Education (ABLE) disclaims any liability with regards to safety in connection with the use of the exercises in its proceedings volumes.

© 1999 Davidson College

Contents

Introduction.....102
Materials 103
Notes for the Instructor 105
Student Outline 106
Acknowledgements.....114
Literature Cited 114
Appendix A: Recipes for solutions 114
Appendix B: Calculation of Allele and Genotype
Frequencies and Hardy-Weinberg Review 115
Appendix C: Prep Schedule 118
Appendix D: Trouble Shooting Cellulose Acetate Gels 119
Appendix E: Alternative Systems for Cellulose Acetate Electrophoresis.....120
Appendix F: Lap Data Sheet.....122

Introduction

Objectives

Many introductory biology students view Hardy-Weinberg Equilibrium as a theory of little or no utility at best, and a form of torture at worst. This attitude is extremely unfortunate since Hardy-Weinberg provides an essential null model for population genetics questions that range from forensic use of DNA profiling to the consequences of habitat fragmentation for endangered species. We designed this lab module to: 1) reinforce information on evolutionary mechanisms already presented in lecture and in the preceding lab, 2) demonstrate the utility of the Hardy-Weinberg Equilibrium Theory, 3) reinforce the relationships between genotype and phenotype, DNA and protein, 4) encourage students to ask questions and construct hypotheses, and 5) introduce students to hypothesis testing and spread sheet software.

Module Description

This lab module was developed for a first year introductory course for prospective science majors that covers evolution, ecology, and physiology. The laboratory allows students to use the Hardy-Weinberg Theory and its assumptions as a research tool. We provide students with collections of seedlings from several populations of the perennial herb white campion (*Silene latifolia*), and ask them to formulate questions regarding the population genetics of these collections (e.g., Do small populations deviate more from Hardy-Weinberg than large populations?). They then use cellulose acetate protein electrophoresis to infer the genotypes of individuals from these populations for the leucine amino peptidase (*Lap*) locus. Since cellulose acetate gels can be purchased pre-made, and since the gels run in 15 min, each group of students can genotype forty to fifty plants in one 3 hr laboratory session. The week following data collection, students use Chi-square tests to compare the observed genotype frequencies with those predicted by Hardy-Weinberg. Students are required to report their results and conclusions in a formal laboratory report written in manuscript form.

This lab could be adapted for upper division courses in evolution or population ecology by incorporating additional loci into the investigations and concepts such as F statistics for characterization of population subdivision. Although this lab uses white campion as its subject, the cellulose acetate technology can be used with many other plant and animal species and many

loci (a search of Agricola or Biosis using the term "cellulose acetate electrophoresis" will help identify protocols for other species).

Module Placement in Laboratory Schedule

This exercise follows a lab segment that uses simulation models to: 1) identify selective and non-selective mechanisms of evolutionary change, and 2) explore the effects of these forces on the genetic composition of simulated populations. Students compare the results of their simulations with Hardy-Weinberg predictions and discuss how different evolutionary mechanisms affect the nature of deviations between simulated results and those predicted by Hardy-Weinberg. We use Bioquest's Evolve software for these simulations because it allows easy capture and manipulation of simulation results. Populus software is suitable for these exercises, but you cannot save the results of simulations and Populus requires IBM compatible computers or MacIntoshes with DOS emulators. Students conduct simulations one week and then report on the results of those simulations and their conclusions the following week. We then use the remainder of the lab period during this second week for discussion of the questions students want to address in the white campion population genetics lab. The evolutionary mechanism simulation lab can be found on the Davidson College Biology Webserver (www.bio.davidson.edu) by selecting the Bio 112 Lab Manual option from the Course Related Materials section.

Timetable

Week 1: Cellulose acetate electrophoresis of Lap locus in white campion - Students use cellulose acetate electrophoresis to collect data on Lap genotypes. Each group of 4 students can collect genotype data for 30 to 50 plants in one 3-hr laboratory session. We give students a take-home quiz on the material provided in the lab instructions the week before this session to insure that they arrive prepared.

Week 2: Analysis of electrophoresis data - Students use Excel to conduct Chi-square tests that compare their observed results with those predicted by Hardy-Weinberg. This session includes an introductory presentation on hypothesis testing and the Chi-square test.

Materials

The quantities specified should cover 4 lab sections with 4 groups/section (4 students/group).

Cellulose acetate electrophoresis equipment

With the exception of the power supplies, Helena Laboratories is the exclusive vendor for the cellulose acetate electrophoresis equipment. Helena Labs does provide educational discounts, but you must ask for this discount. Address: Helena Laboratories, P.O. Box 752, Beaumont, TX 77704-0752. Phone: 1-800-231-5663.

4 Zip Zone Chambers - Helena Cat. No. 1283 (Note: You can make do with only 2 chambers)
4-8 Super Z -12 applicators - Helena Cat. No. 4090
8 Super Z - 12 sample well plates - Helena Cat. No. 4096
4 Super CK aligning base Helena Cat. No. 4094

Note: You can purchase a Super Z- 12 Kit that contains an applicator, a sample well plate, and an aligning base. Helena Cat. No. 4093

1 box of 100 cellulose acetate gels - Titan III - Helena Cat. No. 3033

4 - power supplies capable of supplying 250 volts (You can make do with only 2 power supplies; we use EC 105 power supplies purchased from Fisher)

Grinding and loading supplies

2-mercaptoethanol (5 - 10 ml per lab section)
8 - 16 porcelain well plates (spot plates)
Pasteur pipettes and bulbs
150 small culture tubes (13 X 100 mm) with screw caps to use as grinders - Pyrex No. 9825-13.
Note: This model is available from Fisher. Substitutes should be of comparable size and should have rounded bottoms.
32 small round plastic flower pots (10 cm diameter) to hold tubes
16 shallow containers (at least 10.5 cm square or 11 cm diameter) to hold pots and water (cottage cheese containers work well)
fine sand for grinding
4 microspatulas to use with sand
8 pans to use with ice for grinding (minimum size 15 X 20 cm) - Rubbermaid™ drawer organizers work well
large ice chest filled with ice
8 Sharpie markers - fine point black - not ultra fine point
4 shallow containers (approximately 15 X 20 cm) for soaking gels in buffer solution (Rubbermaid™ drawer organizers or food storage containers work well)
4 automatic pipettes that will deliver a small volume (approximately 10 µl volume of sample with yellow tips; 40 per lab group; 160 per lab section of 16 students)
paper towels - absorbent sheets, not rolls

Staining supplies

8-16 shallow containers (at least 10.5 cm square) to hold gels to be stained plastic wrap
latex gloves (20 pairs / lab section)
50 glass (not plastic) scintillation vials
4 oven timers (to time 15 min and 5 min intervals)
1 water bath to maintain agar at 60-70°C
ring weight to keep agar flask upright in water bath
pipette with automatic extractor or syringe with tubing to deliver 8 ml agar
8 graduated cylinders - 10 ml volume
diaper cloth to cover bench near sink to drain rinsed scintillation vials

Miscellaneous Equipment

Drying oven set at 40-45°C - This allows quick drying of equipment between lab sections.
Hazardous waste containers (for disposal of stained gels and gloves)
Carboy of distilled water - place over sink for rinsing equipment
Box fan to exhaust 2 mercaptoethanol fumes out of window (necessary only if a fume hood is not available)

Solutions/Chemicals (Appendix A provides recipes for solutions)

agar - 1% solution (1 g agar per 100 ml distilled H_2O; boil to dissolve agar) - 250 ml per lab section - Keep at 60-70°C during lab.

grinding buffer (S+) - 100 ml in 10-ml aliquots Make ahead and freeze until needed. You need four 10-ml aliquots per lab section.
Tris-glycine electrode (running) buffer - 4 liters
Phosphate staining buffer - 100 ml per lab section - divide into four small containers (one per lab group)
Fast black K - 25 g
2-mercaptoethanol
Lap substrate solution -Place in four eye dropper bottles (40 ml per lab section).

Notes for the Instructor

Seeds

White campion ranges from Ontario and Quebec south to the valleys of the southern Appalachian Mountains. It grows in hay fields, cow pastures, agricultural fields, and in disturbed habitats such as roadsides. *Silene latifolia* is the currently accepted Latin name for white campion. However, some recent literature refers to this species as *Silene alba*, and older literature refers to it as *Lychnis alba* or *Melindrum*. Seeds mature in mid to late summer depending upon the latitude. Each fruit produces over 200 seeds, and 80 to 90 percent of seeds germinate. Seeds from the populations in southwestern Virginia require one to two months after ripening, stored dry at ambient lab temperatures, for optimal germination. Germinate at 22°C with a 14 hr light/10 hr dark photoperiod. Collect seeds from populations that vary in size and in their proximity to each other.

Preparation of Cellulose Acetate Gels

Soak the gels in the Tris-Glycine electrode (running buffer) for a minimum of 20 min (longer periods of soaking such as 1 to 2 d are fine). If you simply drop the gel into the electrode buffer bubbles will form between the coating and the plastic backing. These bubbles can interfere with conduction of the electrical current and movement of the proteins. To avoid this problem, you need to: 1) grasp the gel between the thumb and forefinger of one hand, 2) place the lower few mm of the gel into the electrode buffer, and 3) then, angle the top of the gel (the edge you are holding) down slowly until the gel is entirely submerged. This method allows the buffer to move up the gel in a single front and greatly reduces the chances of bubble formation. An alternative method involves filling a beaker that is wider and taller than the gel with electrode buffer. Hold the gel vertically and slowly lower it into the beaker.

Lap Substrate

The Lap substrate solution is light sensitive and exposure to high intensities of light can cause the Lap to precipitate. Do not add the Lap substrate to the phosphate buffer until ready to pour the solution over the gel. Conduct the staining in an area free from intense light.

Preparing Students for this Lab

This project involves many steps and students who come to lab unprepared will have difficulty obtaining interpretable results. A take home quiz distributed the week before the lab forces students to read the outline and to understand both the purpose of the study and the purpose of each set of procedures (e.g, extraction). The quiz also allows them to become

familiar with how one scores a gel. We have also found that it pays to have students stop as they move from one procedure to the next (i.e., as they move from extraction to loading the gel) and ask an instructor to check them out on the next procedure.

Appendix B provides instructions on the calculation of allele and genotype frequencies and a review of the Hardy-Weinberg Equilibrium Theory.

Prep Schedule

Appendix C provides a detailed prep schedule.

Trouble Shooting

Appendix D provides a diagnostic chart for problems students commonly encounter with cellulose acetate electrophoresis.

Alternative Systems

Appendix E provides ideas for systems other than white campion that could be used for a population genetics lab.

Student Outline

Background

How do biologists study evolution in natural populations? They employ several approaches; we will use one of these approaches, protein electrophoresis, this week.

Remember that *genes* code for the *amino acid sequences* of *proteins*, and many of these proteins function as *enzymes* or serve as *transporters, structural components, cell recognition factors,* or *hormones*. A *diploid* organism carries two copies of each gene, one that derived from its mother, the other from the father. If the two copies of a gene (*alleles*) differ in their nucleotide sequence, this variation may result in the production of two proteins that differ in their amino acid sequences. These differences in amino acid sequence can translate into differences in the mass and/or charge of the proteins specified by the two alleles. Differences in protein mass or charge can cause the two proteins to migrate at different rates in an electric field. As such, we can use an electric field to separate an individual's proteins, and then stain for the particular protein that interests us. This process is called *protein electrophoresis* and is widely used by evolutionary ecologists and population geneticists to study genetic variation and genetic mechanisms in natural populations.

If a population contains more than one allele for a particular protein, then the proteins that correspond to each allele may move at different rates in the electrical field. Let us consider a population that contains two alleles at a *locus* for a particular enzyme. Some individuals will be *homozygous* for the allele that codes for the more rapidly migrating version of the protein, and their protein will stain as a single band that travels further in the electrical field than the protein of individuals who are homozygous for the other allele (Fig 5.1). Individuals who are *heterozygous* at this locus will display two protein bands (Fig 5.1).

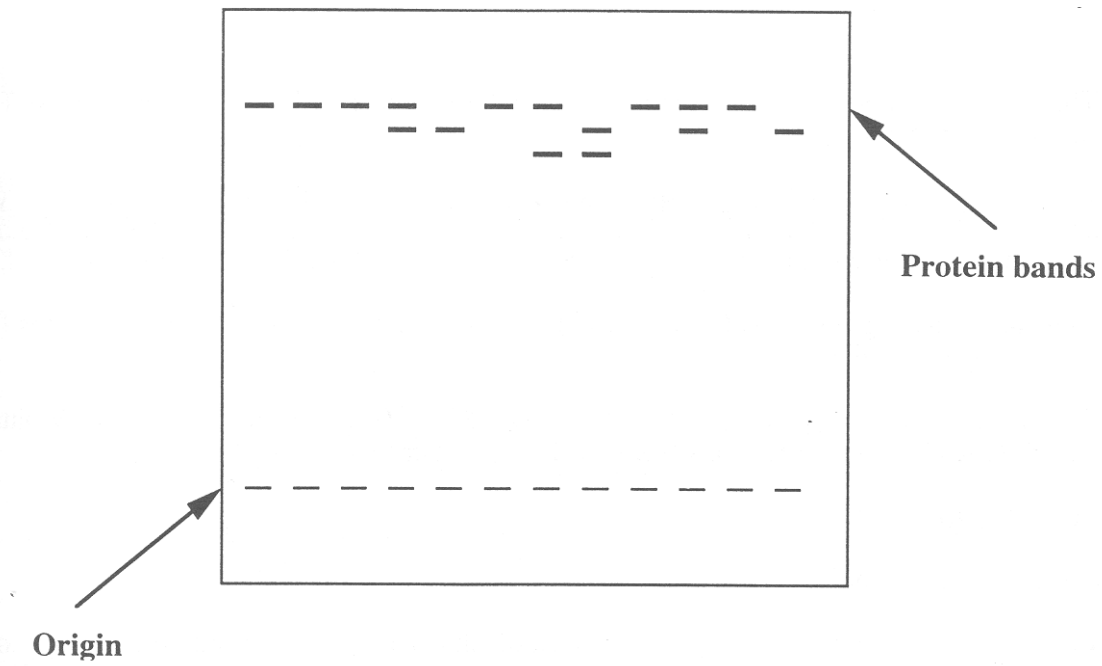

Figure 5.1. Diagram of a cellulose acetate electrophoresis gel for *Lap* in white campion. Each lane represents protein from one individual. Homozygotes display only one protein band, while heterozygotes display two bands. Protein bands are numbered based on their rates of migration, with the fastest migrating band designated as 1. From left to right, the phenotypes are: 11, 11, 11, 12, 22, 11, 13, 23, 11, 12, 11, 22.

The genes we investigate using this technique usually fall into the category of metabolic enzymes, and most of them are involved in cellular respiration or biosynthesis. In most cases we assume that the different alleles present at a locus are selectively neutral, i.e., we assume that the differences in protein structure that result from the presence of more than one allele at a particular locus do not translate into fitness differences among individuals. This assumption is based upon the observation that 1) the *polymorphisms* we find at these loci are common in natural populations and, 2) in populations where we know that *non selective evolutionary mechanisms* are either absent or minimal in their effects, the genotype frequencies for these loci remain in *Hardy-Weinberg Equilibrium*. As such, protein electrophoresis of the proteins specified by these loci can provide us with valuable information on *mating patterns* within populations, *genetic drift*, *founder effects*, and *gene flow* among populations.

In this lab segment, you will ask questions about genotype and allele frequencies for the locus that codes for leucyl amino peptidase (*Lap*), an enzyme that cleaves peptide bonds between leucine and other amino acids, in the plant white campion.

White Campion

White campion (*Silene latifolia*; referred to as *Silene alba* in earlier literature) is a perennial weed in the carnation family. It is native to Europe and was introduced to North

America during colonial times, where it has become naturalized in the northeastern part of the continent. Plants are either male or female, and sex is determined chromosomally in a manner similar to sex determination in mammals. The flowers are pollinated primarily by bees and moths. The seeds are about the size of the poppy seeds you find on your bagel, and we know that dispersal is extremely limited (McCauley, Steven, Peroni, and Raveill, 1996).

The following collections of seedlings or adults plants will be available to you for investigation during this lab.

Whittaker Population - A large population (> 500 plants) located in Eggleston, Giles Co., VA.; plants from the 1992, 1993, 1994, 1995 seed crops. We also have sets of plants from the 1993 seed crop that represent the portion of that cohort that survived 2.5 yr of burial in the soil.

Duncan - A large population (> 500 plants) located ~1 km from Whittaker. Seedlings from the 1994 seed crop are available.

Trailer - A small population (<50 plants), located ~ 10 km from Duncan and Whittaker populations in Giles Co., Va. Seedlings from the 1993 seed crop are available.

Before You Come To Lab

Before lab, each group should:

1. Formulate a populations genetics question that the group can address using protein electrophoresis of the *Lap* locus for one or more of the white campion populations listed above. For example, your group might ask if non random mating occurs in one of these populations while another group could ask if populations from the same general vicinity experience considerable gene flow or operate as a discrete *gene pools*.

2. Establish research and null *hypotheses*. Hypotheses make predictions about your findings, and the null hypotheses always predicts that no real differences exist among groups or between observed results and those predicted by theories. For example, if your group asks if random mating occurs in a population, the your hypotheses would be as follows:

Null Hypothesis - *Lap* genotype frequencies in this population match those predicted by the Hardy-Weinberg Equilibrium theory.
Research Hypothesis - *Lap* genotype frequencies in this population deviate from those predicted by the Hardy-Weinberg Equilibrium theory.

Specify the types of results that would lead your group to reject or accept its null hypothesis.

3. Determine the *research design* for the investigation. This process will include decisions on:
Sample size - A minimum of forty individuals per population is best for this type of investigation. Note: Given time and equipment constraints, if you want to compare two or more populations, join forces with another lab group.

Sample selection - Will you pick the plants you want to use haphazardly, systematically, or randomly?
Sample processing - Will you run all the individuals from one population before you run the individuals from the other population?

Week 1 - Data Collection

This week, you will actually perform cellulose acetate electrophoresis and stain for the *Lap* enzyme. You will use the data you collect to test your hypotheses *Note: the equipment we use for this procedure is very expensive and rather delicate. Please treat it with respect. Students will be billed for equipment damaged due to carelessness.*

Protein electrophoresis includes 5 procedures:

1. Extraction of enzymes from the tissues (grinding).
2. Loading the samples onto the gel.
3. Running the gel (separation of the enzymes in the electrical field).
4. Staining for the enzyme so we can visualize any polymorphisms for the protein.
5. Determining the *Lap* genotypes of individuals based on their electrophoresis phenotypes (scoring the gel).

Extraction

1. Obtain two shallow, rectangular pans. Fill one with ice and place the empty pan on top of the ice filled pan.
2. Obtain a ceramic spot plate and label the wells 1-12 with a Sharpie marker. Place the spot plates into the empty pan that sits atop the ice filled pan.
3. Place a small piece of leaf tissue (approximately 0.5 cm^2 , about the size of the fingernail on your pinky finger) from an individual plant into each well. On your data sheet, identify the plant that corresponds with each well number. Place a leaf sample from *marker plants* A and B into wells number 6 and 7 respectively. Markers are plants whose genotypes have been confirmed by conducting controlled crosses. As such, these plants serve as references that will insure that our gels ran properly and will aid us in the interpretation of our gels.
4. Sprinkle 3 to 5 grains of sand into each well.
5. Obtain a vial of extraction buffer (labelled S+). Place 5 drops of 2-mercapto ethanol into the vial and swirl it gently. Fill a small plastic cup with ice and place the vial of extraction buffer in the ice.
6. Obtain a Pasteur pipette and place 4 drops of the extraction buffer into each well.
7. Obtain a flower pot with culture tubes (grinders) and an empty pot that sits in a shallow container. Fill the container half full with tap water, and place the empty pot in the container.
8. To grind each sample (well), take a clean culture tube, and use the bottom of the tube to grind the contents of the well. Grind until the contents of the well becomes a thin green soup with no visible plant parts. When you finish grinding a sample, place the used culture tube into the flower pot that sits in the container of tap water.

Loading the gel

1. Keep your extracts (ground tissue samples) on ice. Keeping your samples cold is essential to prevent denaturation of your protein or its destruction by proteases.
2. Obtain a Super Z well plate. Using an automatic pipette, take a small sample (10 to 20 ul) of the extract from each spot dish well and place it into one of the small wells on the Super Z well plate. Change pipette tips between samples (wells).
3. When you have filled the wells in your Super Z well plate you are ready to load the gel.
4. Obtain a cellulose acetate gel that was soaked in electrode buffer for at least 20 min. Gently blot the gel dry with a paper towel to remove all surface moisture.
5. Place the gel onto the aligning base, *with the shiny plastic side of the gel down.*
6. Insert the applicator (the gadget with the thin metal tines) into the well plate. Gently press the button top of the applicator 3 times so that the tines on the applicator pick up the samples from the wells.
7. Remove the applicator from the well plate and insert it into the aligning base. Press the button on top of the applicator down onto the gel. Then, while holding the button down, run your index finger *lightly* over the keys. This process loads the samples onto the gel.

Running the gel

1. Hold the loaded gel by its edges and take it to the electrode chamber.
2. Place the gel *coated side down (shiny plastic side up)* in the electrode chamber, with the origin (the end of the gel with the samples) at the negative side of the chamber. Consult Fig. 2.
3. Place a glass slide over each edge of the gel where it touches the paper wicks.
4. Place the cover on the electrode chamber.
5. Attach the chamber leads to the power supply. Turn the power supply on and adjust to 200 volts.
6. Run the gel for 15 min. (Set an oven timer). After 15 min, *turn off the power supply* and then, carefully remove the cover from the electrode chamber. If another gel is running in the same chamber, replace the top and turn the power supply back on.

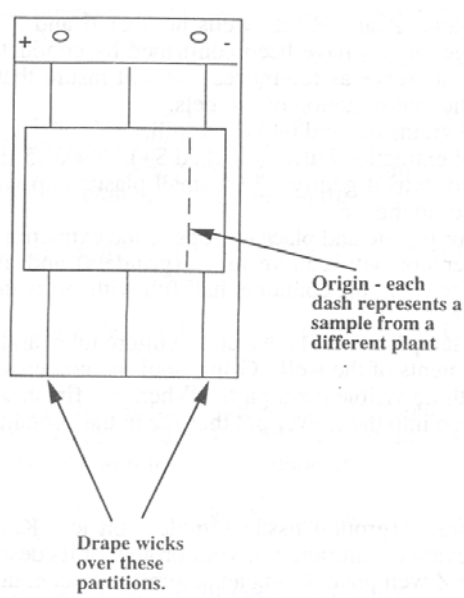

Figure 4.2. Diagram of a cellulose acetate electrophoresis chamber. Pour electrode (running) buffer (Tris-Glycine) into each of the four compartments, but do not let buffer

overflow from one compartment into another. Drape buffer soaked wicks (paper towels work well) over the two outermost partitions. Place gel, stamped side down over wicks, with the origin at the negative pole. Place microscope slides over the gel to ensure good contact with the wicks. *Turn the power supply off before placing your hands in the uncovered chamber*. *Staining the gel*

1. Line a gel box with plastic wrap and place the gel, *coated side up (plastic side down)* into the box.
2. *Put on latex gloves.*
3. Fill a small vial with 5 ml of phosphate buffer. and add 2 droppers-full of *Lap* substrate (approximately 1.5 ml). Swirl the mixture and gently pour it over the gel. Cover the gel for 5 - 10 min. *Note: Lap substrate is carcinogenic and photo-sensitive. Do not add the Lap substrate while standing near bright lights.*
4. In the meantime, fill a small vial with 5 ml of distilled water and add enough fast black K stain to make a mixture that looks like iced tea. Swirl the vial gently to dissolve the stain.
5. After the gel has incubated with the *Lap* substrate for 5-10 min, remove the cover. Take the vial with the fast black K solution and quickly add 5 ml of agar (the agar should be at approximately 60°C). Swirl the vial gently and then pour its contents over the gel.
6. Within 5 - 10 min, bands will appear on the gel. These bands show the presence of the *Lap* protein on the gel.

Scoring the gel

Each lane on the gel represents the *Lap* protein that came from one individual. In white campion, we have identified three alleles at the *Lap* locus. We label proteins (and the alleles they represent) in relationship to their relative rates of migration in the electrical field. Since we find 3 alleles at the *Lap* locus in most white campion populations, the most rapidly migrating *Lap* protein is labelled 1 while the most slowly migrating protein is designated as 3.

Use the sample *Lap* gel provided in Fig. 1 to interpret your gel. Record each individual's genotype on the data sheet (Appendix F) using the 1, 2, 3 designations. For example, an individual that is heterozygous for the most rapidly and most slowly moving proteins would have a 1,3 genotype. *Marker plant A* is a 1,3 heterozygote, and *marker plant B* is a 1,2 heterozygote.

Homework

Calculate the allele and genotype frequencies for each population you investigated (*omit data from the marker plants*). Using your allele frequencies calculate the genotype frequencies predicted by Hardy-Weinberg for a population where no evolutionary mechanisms operate. *Bring your calculations to lab next week.* At that time we will use the Chi-square statistical test to evaluate your hypotheses.

For a locus with three alleles such as the white campion *Lap* locus, the Hardy-Weinberg

Equilibrium theory predicts that the genotype frequencies should conform to the following expectations:

Let: frequency allele 1 = p
frequency allele 2 = q
frequency allele 3 = r

Hardy-Weinberg predictions:
frequency 1,1 genotype = p^2
frequency 2,2 genotype = q^2
frequency 3,3 genotype = r^2
frequency 1,2 genotype = 2pq
frequency 1,3 genotype = 2pr
frequency 2,3 genotype = 2qr

Week 2 - Data Analysis

Graphic Representation of Data - Use DeltaGraph to prepare a figure that compares your observed and predicted results (e.g., your observed genotype frequencies with the Hardy-Weinberg predictions).

Hypothesis Testing - For virtually every group, the observed genotypes will differ from the Hardy-Weinberg predictions. What factors could contribute to these discrepancies?

1. Biased sampling
2. Poor methodology or scoring of gels
3. Operation of evolutionary mechanisms in your population (*research hypothesis*)
4. Chance (*null hypothesis*)

Careful planning and attention to detail can minimize the possibility that the first two factors contribute to differences between our observed and predicted values. As such, when we analyze our data, we must determine if discrepancies between observed and predicted variables represent deviations of our population from Hardy-Weinberg assumptions or simply the effects of chance.

We use *inferential statistics* to determine the probability that the deviations of our observed values from the theoretical predictions could result from chance. *If it is very likely that a sample's deviation from Hardy-Weinberg predictions resulted from chance alone, then we cannot* ***reject*** *our null hypothesis. In other words, we will only reject our null hypothesis in favor of our research hypothesis in cases were the probability that our deviations from Hardy-Weinberg result from chance alone are very low.* How low is low? *We only reject our null hypothesis in cases where the probability that the deviation between our observed and predicted values results from chance is $\leq$ 0.05.*

So, how do we determine the probability that our sample's deviations from Hardy-Weinberg predictions are due to chance? We calculate a *test statistic* which expresses the magnitude of the differences between our observed and predicted values. For variables such as genotype frequencies we use the *Chi-square (X^2) test statistic*. We calculate our X^2 test statistic using the following formula:

$$X^2 = \sum[(o_i - e_i)^2 / e_i]$$

Where: o_i = the number observed for genotype category i
e_i = the number expected for genotype category i, based

on Hardy-Weinberg predictions

$\sum$ = summation (the equation instructs you to calculate $(o_i - e_i)^2 / e_i$ for each genotype category, and then sum these values for all the genotypes)

Now, let us examine the equation for X^2 carefully. If our observations exactly match Hardy-Weinberg expectations, then X^2 will equal zero. But, if our observations differ greatly from Hardy-Weinberg expectations, then X^2 will be a large value. How large must X^2 be in order for us to reject our null hypothesis? X^2 must be sufficiently large enough so there is < 0.05 chance that we would get such a deviation of observed and expected values due to chance alone. How do we determine the *probability (p)* that any particular X^2 value resulted from chance? We can use a published *X^2 table* or instruct a spreadsheet or statistics software package to calculate the probability for us. In either case, we must calculate the *degrees of freedom* (abbreviated as *df* or *v*) associated with our sample. The degrees of freedom equals the number of categories (in our case genotypes) minus the number of pieces of information in our data set that we used to calculate our expected values. In our case, we used the sample size and our estimates of the frequencies of two of the alleles in our populations in order to calculate the number of individuals of each genotype predicted by Hardy-Weinberg (once we estimated frequencies for two of the alleles we could estimate the frequency of the third allele by subtraction). *As such, our degrees of freedom = 6 genotypes - 3 = 3.*

We will use Excel, a spreadsheet software package to calculate X^2. Consult Figure 3 for a sample Excel data sheet. We will use the CHIDIST function on Excel to determine the probability that any particular X^2 value resulted from chance. To do so type:

= CHIDIST(X^2,df)

Excel will return the probability. If the probability is $\leq$ 0.05, then we reject the null hypothesis and conclude that our population probably violates at least one Hardy-Weinberg assumption (i.e., at least one evolutionary mechanism operates on our *Lap* locus in this population).

If the probability (*p value*) associated with our X^2 is > 0.05, then we cannot reject the null hypothesis. We conclude that we do not have enough evidence to argue that evolutionary mechanisms operate on the *Lap* locus in our population.

The instructor will assist groups who want to compare genotype and allele frequencies from two or more populations.

Lap genotype	# of genotype observed (= o)	HW predicted freq	# of genotype predicted by HW (= e)	[(o-e)^2]/e
11	18	0.19	15	0.6
12	26	0.29	23	0.391304348
13	6	0.2011	16	6.25
22	8	0.1111	9	0.111111111
23 & 33	20	0.207	16	1
			X2 =	8.35
			degrees of freedom (df) =	5- 3 =2
			Probability (p) >= Chi square	0.015356634
			Conclusion: Reject null hypothesis; evolutionary mechansim(s) most likley affect the Lap locus in this population	

Figure 5.3. Sample Excel worksheet to test the null hypothesis that a population's true genotype frequencies meet Hardy-Weinberg expectations. Because the sample size was modest, the predicted number of individuals for the rare 33 genotype was < 5. Since the Chi-square test requires that all predicted values be $\geq$ 5, we pooled the data for the 23 and 33 genotypes to create a single category with a predicted value > 5. As such, our degrees of freedom were reduced from 3 to 2.

Acknowledgements

We thank C. R. Werth for showing us how to equip an electrophoresis lab that can accommodate large numbers of users efficiently and economically.

Literature Cited

Hebert, P.D.N. and M.J. Beaton. 1993. Methodologies for Allozyme Analysis Using Cellulose Acetate Electrophoresis: A Practical Handbook. Helena Laboratories, Beaumont, TX.

McCauley, D.E. 1994. Contrasting the distribution of chloroplast DNA and allozyme polymorphisms among local populations of *Silene alba*: Implications for the study of gene flow in plants. Proceedings of the National Academy of Sciences, USA 91:8127-8131.

McCauley, D.E., J.E. Stevens, P.A. Peroni, and J.A. Raveill. 1996. The spatial distribution of chloroplast DNA and allozyme polymorphisms within a population of *Silene alba* (Caryophyllaceae). American Journal of Botany 83:727-731.

Werth, C.R. 1985. Implementing an isozyme laboratory at a field station. Virginia Journal of Science 36:53-76.

Appendix A *Recipes for Solutions*

Tris-Glycine Buffer (Hebert and Beaton, 1993)

30 g Trizma Base (Tris)
144 g Glycine
Bring to volume of 1 liter with distilled water

Dilute 1:9 Tris-Glycine:distilled water for use (Make 1 liter of this diluted concentration.)

Store in refrigerator until ready for use.

S+ Grinding Buffer (Werth, 1985)

100 ml - 0.2 M Tris solution adjusted to pH 8.0 with HCl
0.5 g sodium (meta) bisulfite
0.05 g EDTA, tetrasodium salt
1 ml - 1 M magnesium chloride soln

adjust to pH 7.5 with HCl

Then, add 5 g of polyvinyl pyrolidine (PVP- MW 40,000)
Stir to dissolve (PVP takes about 20-30 min to dissolve)

Place 10 ml aliquots into scintillation vials, label the top of the vial as “S+” and freeze at 20°C until ready for use.

Lap Substrate (Werth, 1985)

Make a 2.5% solution of L-leucyl-beta- naphthylamide in distilled H_2O and place in an eyedropper bottle (2.5% solution = 2.5 g per 100 ml distilled H_2O). Wrap the bottle in aluminum foil to exclude light and store in the refrigerator until ready to use.
Note: L-leucyl-beta- naphthylamide is a carcinogen. Wear gloves and dispose of stained gels and excess solution in chemical waste containers.

Phosphate Staining Buffer (Hebert and Beaton, 1993)

33.12 g $NaH_2PO_4 \cdot H_20$

10.72 g $Na_2HPO_4 \cdot 7H_2O$

Bring to a volume of 1400 ml with distilled H_2O

Appendix B
Calculation of Allele and Genotype Frequencies & Hardy-Weinberg Equilibrium Theory Review

Background

Population geneticists study frequencies of genotypes and alleles within populations rather than the ratios of phenotypes that Mendelian geneticists use. By comparing these frequencies with those predicted by null models that assume no operation of evolutionary mechanisms within populations, we can draw conclusions regarding the evolutionary forces that may influence individual populations. The Hardy-Weinberg Equilibrium Theory serves as the basic null model for population genetics.

Calculation of Allele and Genotype Frequencies
Allele Frequencies

Consider an individual *locus* and a population of *diploid individuals* where two different alleles, A and a, can be found at that locus. If your population consists of 100 individuals, then that group possesses 200 alleles for this locus (100 individuals X 2 alleles at that locus per individual). The number of A alleles present in that population expressed as a fraction of all the alleles (A or a) at that locus represents the *frequency of the A allele* in the population.

To calculate allele frequencies for populations of diploid organisms:

1. Multiply the number of individuals in the population by 2 to obtain the total number of alleles at that locus.
2. Select one of the alleles for your first set of calculations. For example this example, we will choose the A allele from the example provided above.

   Individuals *homozygous* for the A allele will each possess 2 A alleles. Multiply the number of AA homozygotes by 2 to calculate the number of A alleles that group possesses.

   Heterozygotes will each possess only one A allele. The total number of A alleles in the population will equal:

   (the number of A a heterozygotes) + (2 X the number of AA homozygotes)
3. The frequency of the A allele will equal:

$$\frac{\text{total number of A alleles in the population}}{\text{total number of alleles in population for locus}}$$

4. The frequency of the a allele will equal:

   1 - frequency of the A allele

Genotype Frequencies

Consider the same population, locus, and alleles described above. Genotype frequencies represent the abundance of each genotype within a population as a fraction of the population size. In other words, the frequency of the AA genotype represents the fraction of the population that is *homozygous* for the A allele.

To calculate genotype frequencies for populations of diploid organisms:

1. Determine the number of individuals with each genotype present in the population. In the example used above, you would count the number of individuals with the following genotypes:

   AA
   Aa
   aa

2. To determine the frequency of each genotype, divide the number of individuals with that genotype by the total number of individuals in the population.

   Frequency of AA genotype = # AA individuals / population size

   Frequency of Aa genotype = # Aa individuals / population size

   Frequency of aa genotype = # aa individuals / population size

Important Note - Unless you know that a population meets Hardy-Weinberg Equilibrium assumptions, you must use this procedure to calculate genotype frequencies. If you know that a population meets Hardy-Weinberg expectations, then you can calculate genotype frequencies using allele frequencies and the Hardy-Weinberg equations (see below).

Hardy-Weinberg Equilibrium Theory Review

Assertions of the Theory

The *Hardy-Weinberg Equilibrium Theory* refers to loci within populations that experience no *evolutionary mechanisms* (i.e., forces). For such populations the theory asserts that:

1. Allele and genotype frequencies should remain constant from one generation to the next. That is, no evolution should occur at these loci.
2. Given a certain set of allele frequencies, genotype frequencies should conform to those calculated using basic probability. In a one locus/ two allele system such as the one described above, the genotype frequencies should be as follows:

   Frequency of AA genotype = (frequency of A allele)^2
   Frequency of aa genotype = (frequency of a allele)^2
   Frequency of Aa genotype = 2 X (frequency of A allele) X (frequency of a allele)

If the genotype frequencies obtained from a real population do not agree with those predicted by the Hardy-Weinberg Theory, then population geneticists know that some evolutionary mechanism or mechanisms must operate on the locus of interest. A knowledge of the theory can help them narrow down the possible mechanisms. Then, they can use experiments to determine which potential mechanism or mechanisms operate on the locus. As such, the Hardy-Weinberg Equilibrium Theory serves as an important tool for population geneticists.

Assumptions of the Theory (Evolutionary Mechanisms)

Populations will conform to the Hardy-Weinberg Theory assertions only if no *evolutionary forces* or *mechanisms* influence the loci under consideration. The assumptions that populations must meet in order for the Hardy-Weinberg assertions to hold include:

1. Large population size (i.e., no genetic drift)
2. Random mating
3. No difference in the mutation rates between alleles at the same locus
4. Reproductive isolation from other populations (i.e., no gene flow)
5. No differential survival or reproduction among phenotypes (i.e., no natural selection)

Example

Consider a population of 1000 individuals and the locus and alleles described above. Assume that you have no information on the presence or absence of evolutionary mechanism in this population. You find that the population consists of:

90 individuals homozygous for the A allele (AA genotype)
490 individuals homozygous for the a allele (aa genotype)
420 heterozygotes (Aa genotype)

1. Calculate the genotype and allele frequencies for this locus.
2. Determine if this population meets Hardy-Weinberg assumptions (in other words determine if evolutionary mechanisms operate in this population).

Calculation Allele and Genotype Frequencies - Since you do not know if this population meets Hardy-Weinberg Assumptions, you must calculate *both* the allele and genotype frequencies using the raw data.

Allele Frequencies:

The frequency of the A allele will equal:

$$\frac{\text{total number of A alleles in the population}}{\text{total number of alleles in population for locus}} = \frac{[(90*2) + 420]}{(1000*2)} = 0.30$$

The frequency of the a allele will equal:

1 - 0.03 or:

$$\frac{\text{total number of a alleles in the population}}{\text{total number of alleles in population for locus}} = \frac{[(490*2) + 420]}{(1000*2)} = 0.70$$

Genotype frequencies:

Frequency of AA genotype = # AA individuals / population size = 90/1000 = 0.09
Frequency of Aa genotype = # Aa individuals / population size = 420/1000 = 0.42
Frequency of aa genotype = # aa individuals / population size = 490/1000 = 0.49
Hardy-Weinberg Predictions - If no evolutionary mechanisms operate on this locus, then the Hardy-Weinberg Equilibrium Theory predicts that the genotype frequencies should be as follows:
Frequency of AA genotype = (frequency of A allele)^2 = $(0.3)^2$ = 0.09
Frequency of aa genotype = (frequency of a allele)^2 = $(0.7)^2$ = 0.49
Frequency of Aa genotype = 2 X (frequency of A allele) X (frequency of a allele)
= 2*0.3*0.7 = 0.42

Conclusion - Since the observed genotype frequencies equal those predicted by the Hardy-Weinberg Equilibrium Theory, we conclude that no evolutionary mechanisms operate on this locus in this population (i.e., the population meets the assumptions of the Hardy-Weinberg Theory).

Appendix C
Prep Schedule

One Week Before Lab

Prepare One Tray Per Lab Group With:

4 sets of 12 grinders (culture tubes) and soaking apparatus (flower pot in plastic tub)
Super Z - 12 applicator Kit (applicator, sample well plate, and aligning base)
8 empty scintillation vials
1 scintillation vial filled with fast black K
6 glass microscope slides
4 ceramic spot plates (12 wells per plate)
1 microspatula
paper towels (absorbent sheets not roll)
oven timer
Pasteur pipette and bulb
automatic pipette
80 yellow pipette tips
4 shallow pans (container lids, drawer organizers)
2 - 10-ml graduated cylinders (or vials with 5 ml line marked)
small vial of sand for grinding
sharpie - fine point black (not very fine point)
scissors to cut tissue
hazardous solid waste bags
Dropper bottle with Lap substrate
Vial of S+ grinding buffer

One Per 2 Lab Groups:

electrophoresis chamber
power supply
plastic wrap
box of gloves

One for lab:

Ice chest filled with ice (we'll need a lot)
Drying oven set at 40 ^{o}C (move into lab on cart)
water bath for agar - set at 65°C
300 ml of 1% agar
container of 2-mercaptoethanol
Pasteur pipette and bulb to dispense 2-mercaptoethanol
box fan to exhaust 2-mercaptoethanol out window (necessary only if fume hood is not available)

Carboy of distilled water to rinse glassware

One Day Before Lab

Turn on water bath and set to 65°C
Move drying oven into lab on cart and set to 45°C
Place box fan adjacent to window to exhaust 2-mercaptoethanol fumes (only necessary if a fume hood is not available)
Place diaper cloth under 2-mercaptoethanol container and next to sink

Immediately Prior To Lab (i.e., within 1 hr of lab)

Fill ice chest
Remove S+ grinding buffer from freezer (keep on ice once thawed)
Set up gels to soak in Tris-Glycine electrode (running) buffer
Remove Lap substrate, phosphate staining buffer, and electrode buffer from refrigerator
Set up electrode chambers - follow instructions that accompany chamber (paper towels can be substituted for the expensive wicks specified in the manufacturer's instructions)
Microwave agar to melt
Remove small vial of 2-mercaptoethanol from refrigerator and place in fume hood or in front of fan

Appendix D
Troubleshooting Cellulose Acetate Gels
(adapted from Hebert & Beaton, 1993)

Symptom	Potential Causes (* = most frequently encountered)
1. No staining	Stamped wrong side of gel* Ran gel backwards so bands migrated off the gel* Stained gel on wrong side* Chemicals omitted from stain mixture or extraction buffer Chemicals old and degraded
2. No migration of bands	Power supply unplugged* Leads from electrode tray not fully inserted into power supply* Electrodes (platinum wires) not fully submerged in buffer Electrode wicks (paper towels) arranged improperly Electrode wicks and buffer from + and - portions of electrode chambers not mixed after many runs
3. Enzyme migrates in reverse	Ran gel backwards* Leads from electrode chamber to power supply reversed* Some enzymes have a net positive charge and therefore should be oriented with the origin closest to the positive portion of the electrode chamber
4. Enzyme stains weakly	Extracts too dilute (use more tissue or less extraction buffer)* Insufficient sample loaded onto applicator or applied to gel* Chemicals old and degraded
5. Wobbly bands	Gels did not lie flat during gel run Electrode wicks (paper towels) and buffer from + and - portions of electrode chambers not mixed after many runs
6. Wide fuzzy bands	Gel not sufficiently blotted prior to application of samples* pH of extraction buffer incorrect* Insufficient grinding of sample (vegetation clings to applicator)* Extract applied as wide rather than thin lines Applicator dirty

7. Extra unaccountable Contamination among sample wells*
bands Applicator dirty
Contamination of plant tissue samples with insects (e.g., aphids)

Appendix E
Alternative Systems for a Population Genetics Lab

General Protocols for Cellulose Acetate Electrophoresis

Hebert, P.D. N. And M. J. Beaton. 1993. Methodologies for Allozyme Analysis Using Cellulose Acetate Electrophoresis: A Practical Handbook. Helena Laboratories. Beaumont, TX.

This manual provides basic protocols including extraction, stain, and electrode buffers for a variety of allozyme loci. Most of the protocols work well with *Drosophila* spp, but the authors note that cellulose acetate technology has been used on a variety of animal taxa and several plant taxa. Helena Labs provided us with several free copies of this manual when we made our initial equipment purchase.

Alternative Taxa

Silene vulgaris (Bladder Campion - Caryophyllaceae) - McCauley reports that Pgi and Pgm provide a total of 3 highly polymorphic loci in populations from southwestern Virginia. *Lap* also stains, but is only weakly polymorphic in these populations. Protocols follow those provided by Hebert and Beaton (1993).

Pinus resinosa - (Red Pine - Pineaceae) - Mosseler, Innes, and Roberts (1991: Canadian Journal of Forest Research 21:525-528) used cellulose acetate electrophoresis to examine variation at a number of allozyme loci in red pine from Newfoundland.

Pinus banksiana (Jack Pine - Pineaceae) - Gauthier, Simon, and Bergeron (1992: Canadian Journal of Forest Research 22:1958-1965) used cellulose acetate electrophoresis to assay variation for 11 enzyme systems.

Acer saccharum (Sugar Maple - Aceraceae) - Simon, Payette, and Longpre (1995: Canadian Journal of Forest Research 25:743-752) used cellulose acetate electrophoresis to assay variation for 12 polymorphic loci.

Alternatives to Allozyme Electrophoresis

The initial set up for cellulose acetate electrophoresis of allozymes is costly and may be beyond the resources of many departments. However, you can still conduct interesting population genetic experiments provided you find a good natural system. You need a species that is easily accessible and that possesses at least one discrete character that shows heritable variation and that allows students to infer genotype clearly from phenotype. One locus, two allele systems with codominance or intermediate dominance work best.

Before we adopted the white campion electrophoresis lab at Davidson College, we used the common morning glory (*Ipomoea purpurea*) for a population genetics lab. The distribution of pigmentation in the petals of morning glory flowers is controlled by one locus with two alleles (Ennos and Clegg, 1983: Journal of Heredity 74:247-250). Individuals homozygous for one allele show intense pigmentation (pink or deep blue/violet). Individuals homozygous for the other allele have virtually no petal pigmentation (white flowers), and heterozygotes have pale pink or pale blue pigmentation. We took students to populations in the vicinity of campus and had them sample for the three pigmentation morphs. We could determine if individual populations differed from Hardy-Weinberg predictions and if a collection of populations found in the same vicinity showed similar or different genotype frequencies. Students often found significantly higher frequencies of homozygotes than Hardy-Weinberg predicted. They could then generate hypotheses and conduct further field investigations to test these hypotheses.

Since non random mating could result in over abundances of homozygotes, students could study pollinator behavior in the field to determine if individual pollinators tended to be constant to one flower type. Consult Fineblum and Rausher (1997: Ecology 78:1646-1654) for an interesting investigation of he potential of these pigmentation genes to influence resistance to herbivory.

Appendix F
Lap Data Sheet

Group: _____ Date: _____

Gel #1

Lane #	Population	Genotype
1		
2		
3		
4		
5		
6	Marker A	(1,3)
7	Marker B	(1,2)
8		
9		
10		
11		
12		

Gel #2

Lane #	Population	Genotype
1		
2		
3		
4		
5		
6	Marker A	(1,3)
7	Marker B	(1,2)
8		
9		
10		
11		
12		

Gel #3

Lane #	Population	Genotype
1		
2		
3		
4		
5		
6	Marker A	(1,3)
7	Marker B	(1,2)
8		
9		
10		
11		
12		

Gel #4

Lane #	Population	Genotype
1		
2		
3		
4		
5		
6	Marker A	(1,3)
7	Marker B	(1,2)
8		
9		
10		
11		
12		