POPULATION GENETICS OF THE Lap LOCUS IN
WHITE CAMPION (*Silene latifolia*) POPULATIONS

Developed by Dr. Patricia Peroni

BACKGROUND

How do biologists study evolution in natural populations? They employ several approaches, and 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 1). Individuals who are heterozygous at this locus will display two protein bands (Fig 1).
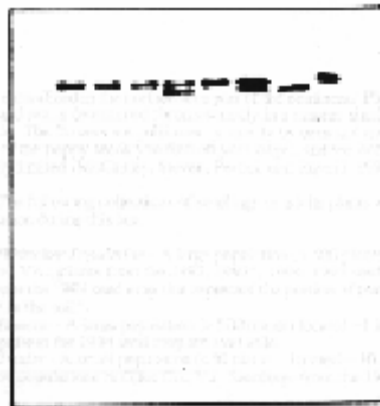

Figure 1: Protein banding patterns for Lap-2 locus in the mollusk *Dreissena polymorpha*.. Each "lane" represents protein from one individual. From left to right, phenotypes are: lanes 1-3 homozygous for the faster migrating protein, lane 4 heterozygous for the fast and slow migrating proteins, lane 5 homozygous for the faster migrating protein, lane 6 heterozygous, lane 7 homozygous for the more slowly migrating protein, lane8 homozygous for the faster migrating protein.

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*) is a perennial weed in the carnation family (Fig 2). 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 et al. 1996).

Figure 2: White campion stems and flowers

The following collections of plants will be available to you for investigation during this lab. All are grown from seeds produced in the summer of 1998.

Whittaker Population - A large population (> 500 plants) located in Eggleston, Giles Co., VA.

Duncan Population - A large population located ~1 km from Whittaker.

Railroad Population – A large population located >10 km from any of the other sampled populations.

Givens Population – A large population located >10 km from any of the other sampled populations.

Beasley Population – A small population (<10 plants) located >10 km from any of the other sampled populations.

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 (40 individuals per population is good 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 (how will you pick the plants you want to use - haphazardly, systematically, or randomly?)
- sample processing (i.e., 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

>Obtain two shallow, rectangular pans. Fill one with ice and place the empty pan on top of the ice filled pan.

>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.

>Place a small piece of leaf tissue (approximately 0.5 cm^2) 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.

Sprinkle a 3 to 5 grains of sand into each well.

Obtain a vial of extraction buffer (labeled S+). Place 5 drops of 2-mercaptoethanol into the vial and swirl it gently. Fill a small plastic cup with ice and place the vial of extraction buffer in the ice.

Obtain a Pasteur pipette and place 4 -5 drops of the extraction buffer into each well.

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.

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

Keep your extracts (ground tissue samples) on ice.

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).

When you have filled the wells in your Super Z well plate you are ready to load the gel.

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.

Place the gel onto the aligning base, with the shiny plastic side of the gel down.

Insert the applicator (the gadget with the thin metal tines) into the well plate. Gently press the button top of the applicator 2-3 times so that the tines on the applicator pick up the samples from the wells.

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

Hold the loaded gel by its edges and take it to the electrode chamber.

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. 3.

Figure 3: Operational electrophoresis tank. Three cellulose acetate plates will fit in one tank simultaneously.

Place a glass slide over each edge of the gel where it touches the paper wicks.

Place the cover on the electrode chamber.

Attach the chamber leads to the power supply. Turn the power supply on and adjust to 200 volts.

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.

Staining the gel

Line a gel box with plastic warp and place the gel, coated side up (plastic side down) into the box.
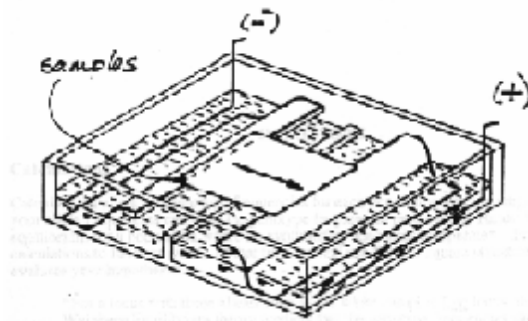
Put on latex gloves.

Fill a small vial with 5 ml of phosphate buffer and add one and one half 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 such as those on your lab bench.

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 ice tea. Swirl the vial gently to dissolve the stain.

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 60oC). Swirl the vial gently and then pour its contents over the gel.

Within 5 - 10 minutes, 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. 4 to interpret your gel. Record each individual's genotype on the data sheet 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
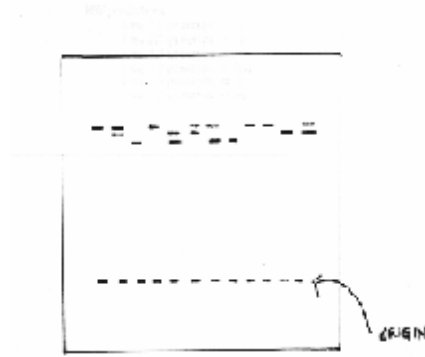genotype. Marker plant A is a 1,3 heterozygote, and marker plant B is a 1,2 heterozygote.

Figure 4: Sample Lap cellulose acetate gel for white campion. From left to right, genotypes are: 11, 12, 33, 11, 23, 12, 13, 33, 11, 11, 22, 12.

HOMEWORK

Calculations

Calculate the allele and genotype frequencies for each population you investigated. Using your allele frequencies calculate the genotype frequencies predicted by Hardy-Weinberg equilibrium 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: freq allele 1 = p
freq allele 2 = q
freq allele 3 = r

HW predictions:
freq 1,1 genotype = p^2
freq 2,2 genotype = q^2
freq 3,3 genotype = r^2
freq 1,2 genotype = 2pq
freq 1,3 genotype = 2pr
freq 2,3 genotype = 2qr

Bio 112　　　　**Lap Lab**　　　　**Data Sheet**　　　　**Davidson College**

Group: _____ Date: _____

Gel #1　　　　　　　　　　　　　　　　Gel #2

Lane #　Population Genotype　　　　　Lane # Population Genotype

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

Gel #3　　　　　　　　　　　　　　　Gel #4

Lane # Population Genotype　　　　　　Lane # Population Genotype

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

WEEK 2

DATA ANALYSIS
Graphic Representation of Data - Use Excel to prepare a figure (bar graph) 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 incases 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 < 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 = 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 calculated the estimated frequencies of two of the three alleles we could determine 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 Fig. 5 for a sample EXCEL data sheet. More detailed instructions regarding the use of EXCEL will be provided in lab. 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:

$= \text{CHIDIST}(X^2, df)$

Excel will return the probability. If the probability is < 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.

I will assist groups who need to compare genotype and allele frequencies for two or more populations.

ACKNOWLEDGEMENTS

Literature Cited

Hartl, D. L. And A.G. Clark. 1989. Principles of Population Genetics. 2nd ed. Sinauer Associates, Sunderland, MA.

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.

Radford, A.E., H.E. Ahles, and C.R. Bell. 1968. Manual of Vascular Flora of the Carolinas. University of North Carolina Press, Chapel Hill, NC.