Allison Smith: South Carolina OBTA Recipient

DNA Profiling Extension Activity:
A study of the blue mussel, *Mytilus*
Teacher Notes

Introduction

In the summer of 2005, I had the opportunity to work with Dr. Thomas J. (Jerry) Hilbish at the University of South Carolina. The research focus of his lab is ecological genetics and the physiological ecology of marine invertebrates. This lesson plan is the direct result of my eight week experience.

The blue mussel, *Mytilus*, is found in intertidal zones in many locations around the world. In California, two species of interest are *M. trossulus* (a native species) and *M. galloprovincialis* (an introduced species). *M. trossulus* prefers cooler temperatures while *M. galloprovincialis* prefers warmer temperatures. The study of the distribution of these two species provides opportunities to learn more about the impact of introduced species on the native species including competitive exclusion in zones of introgression. Also, because of the different temperature preferences, if it is found that the species are spreading towards the north it could possibly be an indicator of changes caused by global warming. Finally, population genetics studies can be conducted by comparing modern and historical data.

In this activity, students learn how to read DNA profiling gels, calculate genotype and allele frequencies, and perform statistical analysis to compare two sets of data.

Related content:
- Mendelian genetics
- Chi-squared test
- Biotechnology
- Evolution
- Population genetics
- Hardy-Weinberg principle

Links to ecology:
- Populations
- Competitive exclusion
- Introduced species
- Global warming

Implementation

It is best to do this lesson after teaching electrophoresis. I usually do it in conjunction with evolution to show population genetics in action. It may be helpful for the students to have had prior exposure to the Chi-squared test and the Hardy-Weinberg equation, but it is not necessary.

I have used this in Biology I classes (college preparatory and honors levels) and in AP and IB HL classes. The difficulty can be increased by not providing a data table and

requiring the students to develop their own using a computer spreadsheet program such as excel, requiring computer generated graphs, and requiring a formal conclusion of the results.

Glu-5' gel basics

- The gels shown have been stained with ethidium bromide and visualized using ultraviolet light. This is why the background appears dark and the DNA bands are bright.

- The difference in band size and brightness is because of different amounts of DNA in the samples (some of the mussels are really small).

- Streaks in the lanes are caused by samples that have large amounts of DNA (some of the mussels are really big).

- The bright band at the bottom of each lane (arrow) is caused by the primers that were used during PCR. It should not be taken into consideration when determining the samples' genotypes.
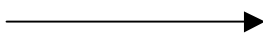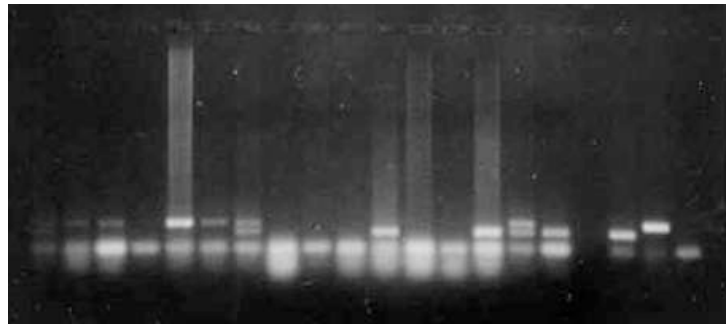
- Lanes 1-16 contain DNA from one *Mytilus* specimen in each. In other words, each lane represents a different organism.

- Lane 17 is empty

- Lane 18 is a sample of *M. galloprovincialis* as a control

- Lane 19 is a sample of *M. trossulus* as a control

- Lane 20 is a blank and contains no DNA. The band that shows is the primers from PCR

- It is helpful to use a ruler to line up the bands with the control bands.

- Some of the lanes may not be readable. These should not be included in the final frequency counts. Also, the students may have some variation in the results due to the individual interpretation of the gels. It is a great opportunity to discuss what scientists would do in these situations.

Answer Key

Practice Gel (Student page 2)

Lane	Genotype
1	GT
2	TT
3	TT
4	-
5	TT
6	TT
7	GT
8	-
9	-
10	-
11	GG
12	-
13	-
14	GG
15	GT
16	GG

Totals:

GG: 3

GT: 3

TT: 4

Total: 10

Frequencies:

GG: 3 GG/10 total organisms= .30

GT: 3 GT/10 total organisms=
.30   G: [2(3 GG) + 3 GT]/(10 × 2)= 9/20 = .45

TT: 4 TT/10 total organisms=   T: [2(4 TT) + 3 GT]/(10 × 2) = 11/20 = .55
.40

Muir Beach 2005 Gels (Student page 3)

Lane	Genotype	Lane	Genotype
1	-	33	GT
2	GT	34	TT
3	GT	35	-
4	TT	36	-
5	-	37	-
6	-	38	GT
7	GG	39	TT
8	GG	40	-
9	TT	41	TT
10	-	42	TT
11	TT	43	-
12	GT	44	-
13	-	45	-
14	-	46	-
15	TT	47	GT
16	-	48	GT
17	TT	49	-
18	TT	50	-
19	TT	51	TT
20	TT	52	TT
21	-	53	-
22	TT	54	-
23	TT	55	TT
24	GT	56	TT
25	-	57	GT
26	-	58	-
27	GT	59	-
28	-	60	-
29	TT	61	TT
30	-	62	TT
31	-	63	TT
32	TT	64	GT

Muir Beach 2005 Data Analysis (Student page 4)

Totals:

GG: 2

GT: 11

TT: 23

Total: 36

Frequencies:

GG: 2 GG/36 total organisms= .05

GT: 11 GT/36 total organisms= .31

TT: 23 TT/36 total organisms= .64

G: [2(2 GG) + 11 GT]/(36 × 2)= 15/72 = .21

T: [2(23 TT) + 11 GT]/(36 × 2) = 57/72 = .79

Muir Beach 1995 Data Analysis (Student page 4)

Number of organisms of each type

GG = .32 × 50 = 16

GT = .16 × 50 = 8

TT = .52 × 50 = 26

Allele frequencies

G = [2(16 GG) + 8 GT]/50 × 2 = .40

T = [2(26 TT) + 8 GT]/50 × 2 = .60

Chi-Square Test (Student page 5)

Genotype	Observed (2005)	Expected (1995)
GG	2.5	16
GT	15.5	8
TT	32	26
Total Number	50	50

$\chi2 = \sum \dfrac{\text{(Observed number - Expected number)}^2}{\text{Expected number}}$

$\chi2 = \dfrac{(2.5-16)^2}{16} + \dfrac{(15.5-8)^2}{8} + \dfrac{(32-26)^2}{26}$

$\chi2 = 11.3 + 7.0 + 1.38$

$\chi2 = 19.68$

degrees of freedom = (3-1) = 2

$p < .01$ Yes, this is significant.

Allele	Observed (2005)	Expected (1995)
G	20.5	40
T	79.5	60
Total number	100	100

$\chi2 = \dfrac{(20.5-40)2}{40} + \dfrac{(79.5-60)2}{60}$

$\chi2 = 9.5 + 6.3$

$\chi2 = 15.8$

d.f. = 1

$p < .01$ Yes, this is significant

Questions (Student page 7)

1. The *Mytilus* population of 1995 consisted of 52% *M. trossulus* individuals. 32% of the population was *M. galloprovincialis* and 16% of the population was hybrids.

2. The *Mytilus* population of 2005 consisted of 64% *M. trossulus* individuals, only 5% of the population was *M. galloprovincialis* and 31% of the population was hybrids.

3. The GG frequency dropped from .32 to .05. The GT frequency increased from .16 to .31 and the TT frequency also increased from .52 to .64.
4. The G allele dropped from .40 to .21 and the T allele increased from .60 to .79.

5. If the chi-square value is not significant, it means that the *Mytilus* population has not changed.

6. If the chi-square value is significant, it means that the *Mytilus* population has changed.

7. There are much fewer M. galloprovincialis and more of the hybrids. The M. trossulus levels have risen some.

Allison Smith: South Carolina OBTA Recipient

DNA Profiling Extension Activity:
A study of the blue mussel, *Mytilus*

Introduction:
DNA profiling has tremendous applications in forensics, paleontology, and determination of paternity to name just a few. It is also a valuable tool in determining what species an organism is. In this activity you are a scientist that is studying two species of blue mussel, *Mytilus trossulus* and *M. galloprovincialis* that are sympatric (found together) along the coast of central California. *M. galloprovincialis* is an introduced species that prefers warmer climates and it appears to be impacting the native *M. trossulus* which prefers cooler temperatures.

The process of preparing the DNA profiles of the *Mytilus* samples has several steps. First, the DNA must be extracted from the organisms. The polymerase chain reaction (PCR) is then performed to amplify the regions of the DNA containing the gene of interest (in this case, the *Glu-5'* gene). The DNA is then digested with restriction enzymes. Finally, the samples are electrophoresed and the gels are analyzed to determine the genotypes of the organisms.

One of the genes that is used to identify species of *Mytilus* is the gene for polyphenolic adhesive protein. This gene, *Glu-5'*, codes for one of the proteins that the mussels secrete in order to attach to a substrate. The gene has two alleles: *M. trossulus* is homozygous for the T allele and *M. galloprovincialis* is homozygous for the G allele. Hybrids will contain both alleles: GT.

Related content:	Links to ecology:
• Mendelian genetics	• Populations
• Chi-square test	• Competitive exclusion
• Biotechnology	• Introduced species
• Evolution	• Global warming
• Population Genetics	

Objectives:
8. Identify the species of *Mytilus* individuals using DNA profiling.
9. Determine the allele and genotype frequencies of *Mytilus* populations in 2005.
10. Compare the 2005 population to the1995 population.
11. Evaluate the genetic changes in the populations.

Procedure:
- Learn how to analyze DNA fingerprinting gels
- Analyze gels to determine the genotypes of the 2005 *Mytilus* population
- Determine the allele and genotype frequencies for the 2005 population
- Perform a chi-square test to compare the 1995 and 2005 populations

How to analyze *Glu-5'* gels

Remember the *Glu-5'* gene has two alleles: *M. trossulus* is homozygous for the T allele and *M. galloprovincialis* is homozygous for the G allele. Hybrids will contain both alleles: GT.

Examine the gel below:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
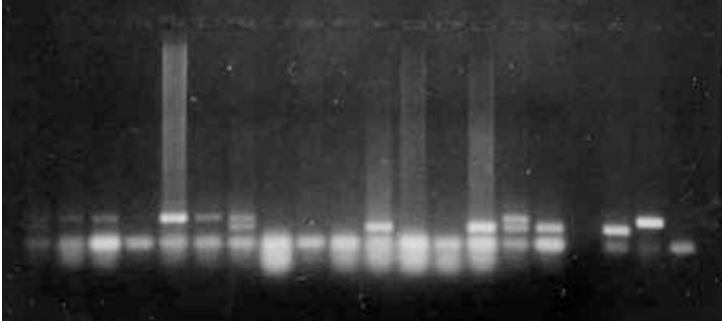
Lanes 1-16 contain DNA of sample organisms

Lane 17 is empty

Lane 18 contains a control of *M. galloprovincialis*

Lane 19 contains a control of *M. trossulus*

Lane 20 contains a blank that has no DNA.

Look at lane 1. If you look carefully, you can see that there are two bands present. Use a ruler to line up the base of the band with the controls. Notice that the bottom band lines up with the *M. galloprovincialis* control and the upper band lines up with the *M. trossulus* control. This indicates that organism #1 is heterozygous (GT) for the Glu-5' gene. Now look at lane 2. Notice there is only one band and it matches the *M. trossulus* control. Therefore, organism #2 is homozygous for the trossulus allele (TT). Now look at lane 4. There is no result that is readable so this organism is not counted in the data analysis. Can you identify the other 13 organisms on this gel?

1.____   5.____   9.____   13.____

2.____   6.____   10.____   14.____

3.____   7.____   11.____   15.____

4.____   8.____   12.____   16.____

Now, calculate the allele and genotype frequencies for this population:
To calculate genotype frequencies, divide the number of organisms of one genotype by the total number of organisms.

Ex: GG = 3/10 = .3

To calculate allele frequencies take the number of GG organisms and multiply by 2. Add the number of GT organisms and divide by the total number of organisms multiplied by 2.

Ex: G = [2(3) + 3]/ (10*2) = 9/20 = .45
How many organisms are GG?_____

How many organisms are GT?_____   Frequency of GG_____ GT _____ TT_____

How many organisms are TT?_____   Frequency of G allele _____

Total number of organisms   _____   Frequency of T allele _____

Muir Beach 2005 Gels

The images below are of gels run on samples of *Mytilus* collected from Muir Beach, CA in 2005. Your task is to determine the genotype of each individual and record it in the data table.

REMEMBER: The last three lanes of each row on a gel contain controls of *M. galloprovincialis*, *M. trossulus*, and a blank. They should not be included in your results.

Lane	Genotype	Lane	Genotype
1		33	
2		34	
3		35	
4		36	
5		37	
6		38	
7		39	
8		40	
9		41	
10		42	
11		43	
12		44	
13		45	
14		46	
15		47	
16		48	
17		49	
18		50	
19		51	
20		52	
21		53	
22		54	
23		55	
24		56	
25		57	
26		58	
27		59	
28		60	
29		61	
30		62	
31		63	
32		64	

Glu-5' Gel 1
Top row is organisms 1-16
Bottom row is organisms 17-32
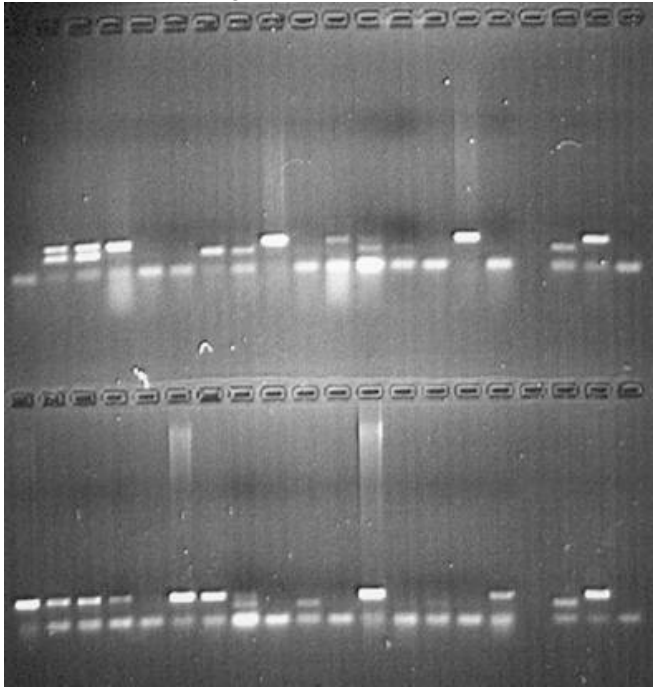
Glu-5' Gel 2
Top row is organisms 33-48
Bottom row is organisms 49-64

Muir Beach 2005
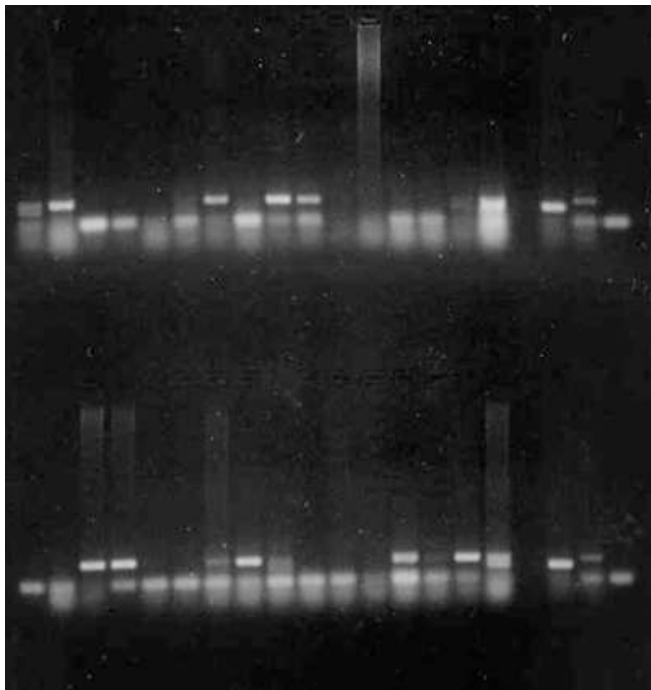
Data Analysis

1. Count the number of samples that had each genotype and record:

       GG _____

       GT_____

       TT_____

       Total _____ (Do not include samples that the genotype was not able to be determined).

2.Calculate the genotype frequencies (show your work) and record:

       GG _____

       GT _____

       TT _____

3.Calculate the allele frequencies (show your work) and record.

       G allele _____

       T allele _____

Comparison of Muir Beach 2005 to 1995

The data below represents the genotype frequencies for Muir Beach in 1995:

      Number of samples: 50   Genotype frequencies:

                                        GG: .32
                                          GT: .16
                                          TT: .52

Determine the number of organisms of each genotype (multiply the genotype frequency by the total number of organisms). Be sure to show your work:

       GG _____

       GT _____

       TT _____

Calculate the allele frequencies for 1995 (show your work):

G _____

T _____

The Chi-Square Test

The chi-square test is a statistical test that helps scientists determine if differences in observed data compared to expected data are due to chance or if they are significant. We can formulate a null hypothesis that the Muir beach mussel population has not changed since 1995 therefore the 2005 frequencies should be the same. The formula for calculating the chi-square value is

$$\chi2 = \sum \frac{(\text{Observed number - Expected number})^2}{\text{Expected number}}$$

Once the chi-square value has been calculated, it is used to determine the probability (p) that the difference in the observed and expected values is due to chance. If $p \le .05$ the difference is considered significant and not due to chance so the null hypothesis is not accepted. In other words, the population has changed since 1995.

Use the chi-square formula to determine if the genotype frequencies of the 2005 population have changed significantly from the 1995 population. When working with chi-square, numbers of organisms must be used, not frequencies. Since the 1995 population contained 50 organisms, we must calculate the observed values for 2005 based on a population of 50 organisms.

- To calculate the observed number for 2005, multiply the genotype frequency by 50. Do this for each genotype and record in the table.

- You have already calculated the number of organisms for 1995 (see p. 5). Transfer these numbers to the table.

Genotype	Observed (2005)	Expected (1995)
GG		
GT		
TT		
Total Number	**50**	**50**

- Use the observed and expected values in the table to calculate the chi-square value (show your work):

$$\chi 2 = \sum \frac{(\underline{\text{Observed number - Expected number}})^2}{\text{Expected number}}$$

$\chi 2 =$ _____

- Determine if the χ2 value is significant

    - Calculate the degrees of freedom (d.f.). This is equal to the number of terms in the chi square equation -1.

    - Find the probability using the chi-square table of distributions. Choose the correct row based on the degree of freedom. Locate the χ2 value and look at the column headers. This is the probability that the difference in the two populations is due to chance. if $p \leq .05$, then the null hypothesis is not accepted and the population has changed significantly since 1995.

        $p =$ _____   Is this significant? _____

The Chi-square Table of Distributions

d.f.	P= 0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01
NOT SIGNIFICANT (spanning P=0.90 through 0.10)								**SIGNIFICANT** (spanning 0.05, 0.01)	
1	0.016	0.064	0.148	0.455	1.07	1.64	2.71	3.84	6.64
2	0.211	0.446	0.713	1.38	2.41	3.22	4.60	5.99	9.20
3	0.548	1.00	1.42	2.37	3.66	4.64	6.25	7.82	11.3

Perform a chi-square test on the allele frequencies for the 1995 and 2005 populations. Remember you must convert the frequencies to numbers of alleles. Show your work.

Allele	Observed (2005)	Expected (1995)
G		
T		
Total Number	**100**	**100**

$\chi 2 =$ _____   d.f. = _____ $p =$ _____   Is this significant? _____

Questions

1. Describe the species makeup of the population in 1995.

2. Describe the species makeup of the population in 2005.

3. How have the genotype frequencies changed since 1995?

4. How have the allele frequencies changed since 1995?

5. If the results of the chi-square test are not significant, what does this mean for the two *Mytilus* species?

6. If the results of the chi-square test are significant, what does this mean for the two *Mytilus* species?

7. How has the Muir Beach population of Mytilus changed since 1995?