Chapter 8

An Introduction to Population Genetics

Matthew E. Andersen

Department of Biological Sciences
University of Nevada, Las Vegas
Las Vegas, Nevada 89154-4004

Matthew Andersen received his B.A. in 1986 from Sonoma State University, California. He was an honors student for his undergraduate work and is now a doctoral student at UNLV. He has worked in biomedical research and as a pathologist's assistant. His Ph.D. research is on the molecular systematics of fish, particularly speckled dace, *Rhinichthys osculus* (Cyprinidae).

Reprinted from: Andersen, M. E. 1993. An introduction to population genetics. Pages 141-152, *in* Tested studies for laboratory teaching, Volume 14 (C. A. Goldman, Editor). Proceedings of the 14th Workshop/Conference of the Association for Biology Laboratory Education (ABLE), 240 pages.

Contents

Introduction

This laboratory introduces the beginning biology student to some of the modern principles of population genetics. Hardy-Weinberg equilibrium and the Hardy-Weinberg equation are introduced. Though populations rarely, if ever, fit this equilibrium, it is valuable for evaluating the direction and rate of allelic shifts in populations. For instance, if the frequency of one allele is shown to be increasing dramatically from one generation to the next, we hypothesize that some other force, such as drift or selection, is acting to disrupt the equilibrium.

For some student groups, it may be valuable to stress that scientists do not view evolution as a belief system or religion. Rather, it is the demonstrable change in allele frequencies with time, though usually in geologic, not human time. Organisms with brief generation times (e.g., yeast or *Drosophila*) can be monitored in the laboratory to demonstrate allelic shifts.

This lab requires a very moderate amount of preparation. Arranging decks of cards with appropriate numbers is the most complicated preparatory step. The lack of complicated equipment and set up allows the instructor and the students to focus on the principles and mathematics involved without distraction.

The entire laboratory can be accomplished within 3 hours or less. This includes a 15–20 minute introductory lecture, explaining the activities, carrying out the activities, and demonstrating the concepts over multiple generations with computer software. Even without the addition of software demonstrations, the principles may often be new or the concepts challenging to ideas students have gained previously in their education, particularly more mature students.

The following **materials** are required per student: 3" × 5" card with a genotype written on it (begin the entire class with a 1:2:1 ratio of homozygous dominant, heterozygous, and homozygous recessive genotypes); die (1); coin (a penny will do); deck of playing cards with all the face cards removed (i.e., JQKA); and population genetics software (e.g., *Time Machine*).

Student Outline

The theory of evolution emphasizes that populations, not individuals, evolve. A *population* is all the members of the same species that live in one locale. Each member of a population is assumed to be free to reproduce with any other member, and when reproduction occurs, the genes of one generation are passed on in the manner described by Mendel's Laws. In a population of sexually reproducing individuals, the various alleles of all the genes in all the individuals make up a

gene pool for the population. The gene pool of a population can be described in terms of *allele frequencies*.

Two independent investigators, G. H. Hardy and W. Weinberg, realized that the binomial equation:

$$p^2 + 2pq + q^2$$

could be used to directly calculate the genotype frequencies of a population. This equation has been termed the "Hardy-Weinberg Equation."

Assume that in any generation, each type of sperm had an equal chance to fertilize each type of egg. Assign the letter *p* to the dominant allele (*A*) and the letter *q* to the recessive allele (*a*); then

	p	*q*
p	p^2	*pq*
q	*pq*	q^2

In this case: p^2 = homozygous dominant individuals (*AA*)
q^2 = homozygous recessive individuals (*aa*)
$2pq$ = heterozygous individuals (*Aa*)

To use this equation to calculate the genotypic frequencies in a population, it is necessary to know the frequencies of each allele. For example, suppose the dominant allele (*p*) has a frequency of 0.7 and the recessive allele (*q*) has a frequency of 0.3. Then the genotype frequencies of the population would be:

p^2 (homozygous dominant) = 0.7×0.7 = 0.49
q^2 (homozygous recessive) = 0.3×0.3 = 0.09
$2pq$ (heterozygous) = $(0.7 \times 0.3) \times 2$ = 0.42
1.00

The frequencies of the two alleles ($p + q$) must equal 1.00 and the frequencies of the various genotypes ($p^2 + 2pq + q^2$) must also equal 1.00.

If you know the percentage of individuals in a population that shows the recessive phenotype, it is possible to calculate the genotype frequencies and the allele frequencies. Suppose, by inspection, you determine 16% of the human population has a continuous hairline. This means that 84% of humans must have widow's peak. Using the Hardy-Weinberg equation, you can determine how many individuals are homozygous dominant and how many are heterozygous.

First, convert 16% to a decimal or frequency: $0.16 = q^2$ and therefore $q = 0.4$. Since $p + q = 1.0$, $p = 0.6$. Now that you have a value for *p* and *q*, you can enter them into the equation and determine:

p^2 (homozygous dominant) = 0.36 or 36%
$2pq$ (heterozygous) = 0.48 or 48%
q^2 (homozygous recessive) = 0.16 or 16%

The Hardy-Weinberg equation can be used to calculate the genotype frequencies of a pair of alleles at a gene locus in an established population. It can also be used to predict genotype frequencies of the next generation as long as there have been no gene pool changes. When there have been no changes, genetic equilibrium will **only** occur when:

1. The population is large enough to be unaffected by random gene changes (i.e., genetic drift).
2. There is no gene flow (immigration or emigration).
3. No mutations occur or there is mutational equilibrium.
4. Reproduction is random (independent of genotype).
5. Natural selection is not acting on a particular phenotype.

In natural populations all of these conditions are almost never met. Why then is the concept of genetic equilibrium such an important concept?

Objectives

1. Define the term gene pool and explain how it relates to the concept of evolution.
2. Understand the concept of genetic equilibrium and the conditions required for it to occur.
3. Explain how genetic equilibrium relates to evolution.
4. Explain how the various simulations deviate from the conditions required for genetic equilibrium.

Since evolution is difficult, if not impossible, to observe in a 3-hour laboratory period, we will simulate the evolutionary process using the class as a reproducing population. We will use five different scenarios to depict changes in genotypic frequencies over time.

Case 1: Random Mating, No Selection

1. Each of you will receive a card indicating your genotype for this experiment. You as a class will become a population of randomly mating individuals. The initial gene frequency of your population is 25% *AA*, 50% *Aa*, and 25% *aa*.
2. Observe your genotype. The allele on the left is designated allele #1; on the right is allele #2.

   To insure random mating you must be completely promiscuous in this experiment. Choose any student (parents need not be of the opposite sex — we are a strange species) in the class and confidently approach them; they cannot refuse. To mate, each parent must contribute one allele to the offspring. To determine which allele is passed on, each parent flips a coin. If the coin comes up heads, then allele #1 is passed on, if tails, then allele #2. You, as a couple, now have produced one offspring. Since we believe in zero population growth (ZPG), each couple must have two children. Repeat the process to produce your second offspring. Record the genotype of both children in the F_1 row in Table 8.1.
3. For the sake of the experiment, we must maintain a constant population size and we do not want one generation breeding with another. *Only breed within the same generation.* Your parent (original) genotype dies and you and your partner become the genotype of your offspring. (Make sure you use *both* genotypes.) When instructed to do so by your instructor, mate again with a *new partner*. Follow the exact same mating procedures and record the offspring in the F_2 row. Continue this process through five generations. Remember to always flip the coin and to change your genotype for each new encounter. Complete Table 8.1 and report your results to your instructor so a class tally can be made.
4. Based on these results did this population exhibit genetic equilibrium? If not which condition was not met?

Table 8.1. Data table for Case 1: Random Mating, No Selection.
Initial (whole population) frequencies: 25% *AA*, 50% *Aa*, 25% *aa*.

Your initial genotype:		
Generation	Child 1	Child 2
F_1		
F_2		
F_3		
F_4		
F_5		
Sum of entire population (class) F_5 offspring		
AA:	*Aa*:	*aa*:
Genotype frequencies		
AA:	*Aa*:	*aa*:
Allele frequencies		
A:	*a*:	

Case 2: Selection

1. Now that you have the facts of life well in hand, we can begin to modify our simulation, making it more realistic. In this case study, we will investigate some basic questions about selection and gene frequencies. In humans there are several genetic "disorders" or mutations that have been well characterized. It is important to remember that selection very rarely acts in such a clear-cut manner. Most genetic mutations have no effect on the organism's fitness, good or bad (Kimura, 1983).

   An example of a mutation with an observable effect on fitness is sickle-cell anemia. This is a single allele trait that produces abnormal hemoglobin. In the days before modern medicine the recessive genotype was lethal; the individual with this genotype was severely anemic and died before reaching reproductive maturity. Both homozygous and heterozygous dominant individuals survive. In this case, we will be selecting against the recessive individuals 100% of the time.

2. The procedure is similar to Case 1. Start again with the *same initial genotype* and choose the genotype of the offspring by flipping your coin. (The population will again begin from an idealized 1:2:1 ratio). This time, however, there is one important difference. **Every time your offspring is *aa* it dies.** Since we want to maintain a constant population size *the same two parents must try again* until they produce a surviving offspring. Only record the surviving offspring in Table 8.2.
3. As you might have realized, there is one real problem that must be solved before we begin. Twenty-five percent of you are already recessive. Instead of killing you off, we will make the assumption that *aa* in the initial parent is not lethal — the mutation begins with your offspring. If you are *aa* originally do not mate with another *aa* in the first round.
4. What happened to the frequency of the recessive allele?

Table 8.2. Data table for Case 2: Selection.
Initial (whole population) frequencies: 25% *AA*, 50% *Aa*, 25% *aa*.

Your initial genotype:		
Generation	Child 1	Child 2
F_1		
F_2		
F_3		
F_4		
F_5		
Sum of entire population (class) F_5 offspring		
AA:	*Aa*:	*aa*:
Genotype frequencies		
AA:	*Aa*:	*aa*:
Allele frequencies		
A:	*a*:	

Case 3: Heterozygote Advantage

1. Case 2 demonstrated that the lethal recessive gene is rapidly decreasing in the population. Scientific studies of actual populations of African Blacks show an unexpectedly high frequency of the sickle-cell allele present in some populations. Our simulation does *not* accurately reflect the real situation. In the real situation malaria plays a role in selection along with sickle-cell anemia. Malaria, a disease caused by a parasitic protozoan, is a prevalent in areas with high frequency of the sickle-cell trait. Malaria is often fatal, but what became apparent after years of investigation, was that heterozygous individuals (*Aa*) did not die of malaria. The recessive sickle-cell hemoglobin trait was not expressed in these individuals sufficiently to cause sickle-cell anemia. It was expressed sufficiently enough, however, to make such individuals undesirable hosts for malaria parasites. Heterozygous individuals suffer neither fatal sickle cell-anemia nor fatal malaria.
2. In this case, we will simulate the population that carries sickle-cell traits and is exposed to malaria. In this round, keep everything the same as Case 2 (*aa* offspring die). In addition, **if you have an offspring that is homozygous dominant (*AA*), it will die of malaria half the time**. To determine if your offspring will survive, roll a die. If you get 1, 2, or 3, the individual lives. If you get 4, 5 or 6, the individual dies of malaria. *Keep mating with the same partner until you have two surviving offspring*. Record each generation in Table 8.3 and complete five generations as in the other exercises.

3. How did the frequency of heterozygous individuals change?

Table 8.3. Data table for Case 3: Heterozygous Advantage.
Initial (whole population) frequencies: 25% *AA*, 50% *Aa*, 25% *aa*.

Your initial genotype:		
Generation	Child 1	Child 2
F_1		
F_2		
F_3		
F_4		
F_5		
Sum of entire population (class) F_5 offspring		
AA:	*Aa*:	*aa*:
Genotype frequencies		
AA:	*Aa*:	*aa*:
Allele frequencies		
A:	*a*:	

Case 4: Genetic Drift

Genetic drift is a change in gene pool variation that occurs purely as a result of *chance*. The randomness of this phenomenon is sometimes referred to as its *stochastic* nature. Genetic drift always influences gene frequencies to some degree, no matter how large the population. In small populations of less than approximately 100 individuals, the alteration of gene frequencies may be quite dramatic, usually causing the loss of one or the other allele entirely. The loss of alleles due to genetic drift may occur much more quickly than losses by natural selection. In general, genetic drift will cause populations to lose their heterozygotes, and to become comprised exclusively of the genotypes homozygous dominant and homozygous recessive (illustrated in Figure 8.1).

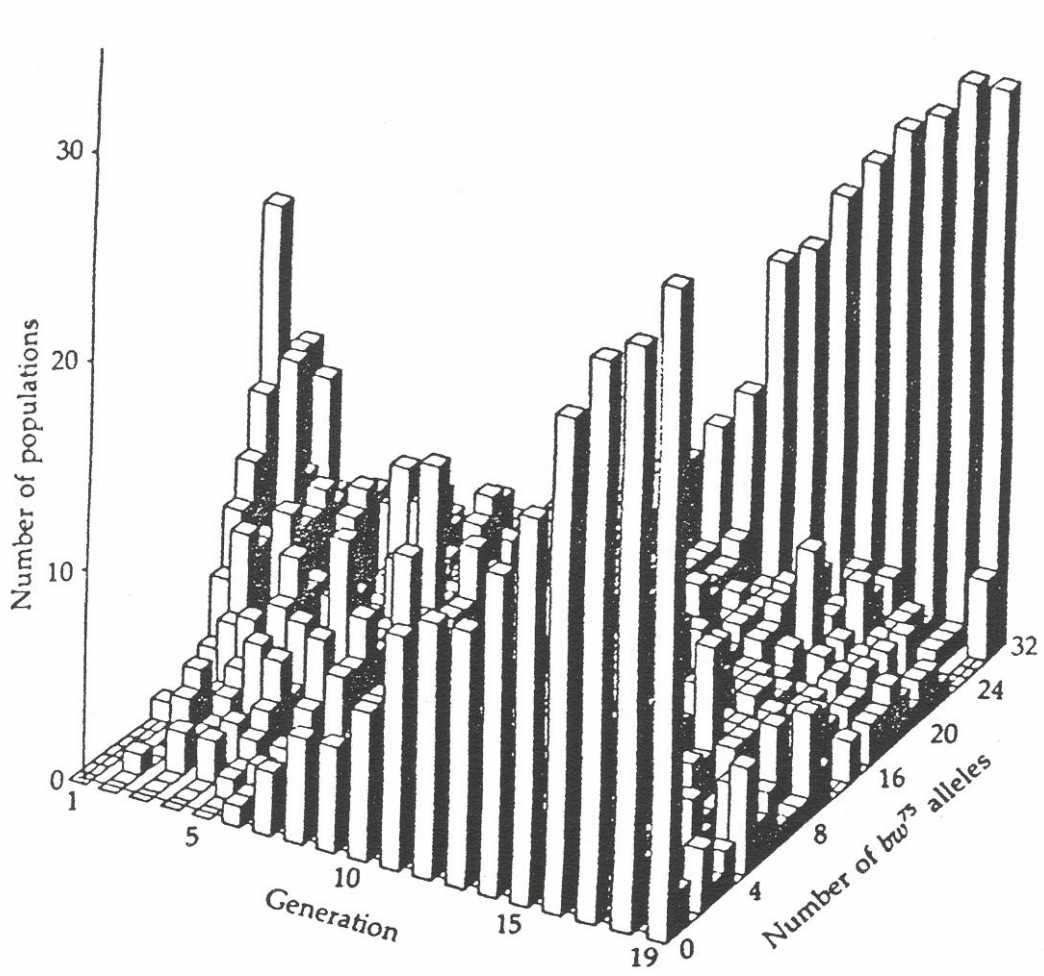

Figure 8.1. A population of fruit flies (*Drosophila melanogaster*) began with a high percentage of heterozygotes and no homozygotes. After 19 generations, most of the heterozygotes were lost, but the two types of homozygotes had greatly increased their frequency in the population. From Hartl and Clark (1989), based on data from Buri (1956).

In its most extreme case, genetic drift leads to loss of certain alleles. It has been observed that genetic drift is particularly evident after a new population is started by only a few members of an originally large population. Imagine that by chance a small number of individuals found a colony; only the founder's alleles are passed on to the next generation. This is known as the *founder effect*, and it can result in an isolated population that has radically different genotypes than the population as a whole. With little variability in the initial population, genetic drift is likely to cause the loss of one allele at each locus.

Populations that are initially large and then are reduced dramatically in size are said to have passed through a *population bottleneck*. For example, this has happened to wild animals many times in the last few centuries as a result of overhunting and habitat destruction. Like the founder effect, the small numbers of genetic variables remaining in the population after passing through the bottleneck make it more likely that genetic drift will influence the composition of the population. When alleles are lost the result is that entire populations are much more susceptible to other

negative phenomenon, such as disease. With sufficient individual variability, organisms may survive many adverse conditions, allowing the entire species to survive.

In this exercise you will simulate genetic drift using a deck of cards.

1. Remove all the face cards from the deck. Shuffle the remaining cards.
2. Deal out 10 cards face up in a row in front of you. These cards represent 10 alleles. Assume that you have started with a frequency of 50% *A* allele and 50% *a* allele. Let cards 1 through 5 represent the *A* allele and cards 6 through 10 represent the *a* allele. Determine the frequency of each allele in your first generation and enter these values in Table 8.4.
3. Pick up the cards, reshuffle and deal out a new set of 10 cards. Using the frequency you calculated for the *A* allele in the first generation, reassign values. For example if you had a 60% frequency for the *A* allele, in this generation, cards 1 through 6 would represent the *A* allele and cards 7 through 10 the *a* allele. Determine the new frequency for the *A* allele.
4. Pick up and shuffle the cards again. Deal out 10 for the third generation and determine the frequencies as in the second generation.
5. Repeat this procedure until one of the alleles appears with a frequency of 100%. Record the number of generations it took for one of the alleles to disappear. Compare with other members of the class.
6. How many generations did it take for one allele in your hypothetical population to disappear? What was the mean number of generations for the class as a whole?

Table 8.4. Data table for Case 4: Genetic Drift.

Generation	Change in frequency	Number of *A* alleles	Number of *a* alleles
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Case 5: Computer Simulations

This exercise involves the use of a computer simulation called *Time Machine*. The major advantage of this simulation over the in-class simulations which you have just completed is the much longer time frame over which the simulations can be run. The in-class simulations were only carried out over five generations, which is a very short period of time in an evolutionary sense. The computer simulations allow you to investigate the same phenomena, but over hundreds or perhaps thousands of generations.

There are three basic simulations which you will conduct: genetic drift, selection, and mutation. Based on your simulations answer the following questions:

Genetic Drift

1. How large must a population be to avoid genetic drift?
2. Starting with a very small population (10), if the simulation is run many times are the results the same? Why?

Selection

1. Does an allele need to be lethal to be selected against?
2. Is it possible to completely remove a *lethal recessive* allele from a population?
3. If possible how many generations would it take to accomplish this?
4. Is it possible to completely remove a *lethal dominant* allele from a population?
5. How many generations would it take to accomplish this?

Mutation

1. At natural mutation rates how does the impact of mutation compare to that of selection?

Notes for the Instructor

The importance of calculating the Hardy-Weinberg equilibrium values is that deviations from the values over generations allow us to quantify evolution by monitoring allele frequencies. In the simulations, be sure that as students conduct their mating experiments they are always mating with students on the same generation. For example, an F_2 offspring should only mate with another F_2 offspring. Staying on the same generation is facilitated by making sure all students understand how the "mating" is accomplished before anyone is allowed to begin. Also ensure that after students mate, each of the offspring is represented in the following generation by having both parents choosing which child they will be. For example, if you and I "mated," producing one *Aa* and one *aa* child, I would be *Aa*, you would be *aa* in the following round of mating, ensuring that all alleles were accurately represented.

In the three mating experiments, I would expect that the random mating would result in allele frequencies very similar to the initial 1:2:1 ratio, although sometimes genetic drift will act and cause allelic shifts, particularly with smaller groups. When the homozygous recessives are eliminated,

they of course are not represented in the final frequencies. When heterozygotes have some selective advantage (Case 3), the frequency of that genotype will increase.

When conducting the genetic drift exercise, some students will fix one allele or the other in less than 10 generations, others will not. One of the more interesting results is when one allele is nearly fixed, and then the tide of chance turns and the opposite allele is finally fixed.

After diligent searching, it appears that BioSoft, the company that originally manufactured the software in this exercise, is no longer in business. I will be indebted to anyone that cannot locate an official distributor of this program. If you cannot locate a source for population genetics software, you can contact me for a copy of *Time Machine*. Send an unformatted or IBM-formatted diskette and return postage to: Matthew Andersen, UNLV – Biological Sciences, 4505 Maryland Parkway, Las Vegas, Las Vegas, Nevada 89154-4004. The background information associated with a population genetics programs will provide the highlights of the software. The objective of using such a program is to illustrate what occurs in nature over many generations, the actual time frame of evolution. Begin the computer illustrations with 5 or 10 generations, as the students do in their exercises, and quickly increase generations by orders of magnitude (i.e., 100s and 1,000s).

Acknowledgments

I am indebted to Bill Wischusen and Roberta Williams for their assistance and guidance on this and many other projects.

Literature Cited

Buri, P. 1956. Gene frequency in small populations of mutant *Drosophila*. Evolution, 10:367–402.

Hartl, D. L., and A. G. Clark. 1989. Principles of population genetics. Second edition. Sinauer Associates, Sunderland, Massachusetts, 682 pages. [ISBN 0-87893-302-6]

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, 367 pages. [ISBN 068509-798-6, 317932]