AP Biology **Name_____**

Lab 2: Mathematical Modeling: Hardy-Weinberg[1]

Overview

In this lab you will:

1. learn about the Hardy-Weinberg law of genetic equilibrium, and
2. study the relationship between evolution and change in allele frequency by using a mathematical model to demonstrate what can happen over many generations

Objectives

Before doing this lab you should understand:

1. how natural selection can alter allelic frequencies in a population
2. the Hardy-Weinberg equation and its use in determining the frequency of alleles in a population, and
3. the effects on allelic frequencies of selection against the homozygous recessive or other genotypes.

After doing this lab you should be able to:

1. calculate the frequencies of alleles and genotypes in the gene pool of a population using the Hardy-Weinberg formula, and
2. discuss natural selection and other causes of microevolution as deviations from the conditions required to maintain Hardy-Weinberg equilibrium,
3. use a data set that reflects a change in the genetic makeup of a population over time and apply mathematical methods and conceptual understandings to investigate the cause(s) and effect(s) of this change,
4. apply mathematical methods to data from a real or simulated population to predict what will happen to the population in the future,
5. evaluated data-based evidence that describes evolutionary changes in the genetic makeup of a population over time,
6. use data from mathematical models based on the Hardy-Weinberg equilibrium to analyze genetic drift and the effect of selection in the evolution of specific populations,
7. justify data from mathematical models based on the Hardy-Weinberg equilibrium to analyze genetic drift and the effect of selection in the evolution of specific populations,
8. describe a model that represents evolution within a population, and
9. evaluate data sets that illustrate evolution as an ongoing process.

Introduction to Hardy-Weinberg Equilibrium

In 1908 G.H. Hardy and W. Weinberg independently suggested a scheme whereby evolution could be viewed as changes in the frequency of alleles in a population of organisms. They argued that if a population met five conditions that there would be no changes in allele frequencies from generation to generation. Those five conditions are as follows:

1. The breeding population is large, reducing the effect of chance on changes in allele frequencies
2. Mating is random with respect to the gene in question. Individuals show no mating preference for a particular phenotype.
3. There is no mutation of the alleles.
4. No individuals enter or leave the population by migration (no immigration or emigration).
5. There is no selection. In other words, the alleles of the gene in question do not affect mating or survival, so every genotype has an equal chance of surviving and reproducing.

[1] Adapted from the College Board AP Biology Student Lab Manual, 2001 edition and 2012 edition, as well as materials from Jon C. Herron (© 2003) and Amherst College

Hardy and Weinberg also described the distribution of alleles and genotypes within a population mathematically. The following variables are used, and each is expressed in decimal form:

p = frequency of the dominant allele
q = frequency of the recessive allele

Note that p + q must always equal 1. The analysis does not work if there are more than 2 alleles of the gene being studied. (It can be done, but the math is much more complicated.)

In a population, p is also the likelihood of a dominant allele being passed on to an offspring, and q is the likelihood of a recessive allele being passed on. Thus, we end up with the following conclusions:

p^2 = frequency of homozygous dominant individuals in the population
$2pq$ = frequency of heterozygous individuals in the population
q^2 = frequency of homozygous recessive individuals in the population

Note that $p^2 + 2pq + q^2$ should always equal 1.

How can this rule be applied to real populations? One cannot see the frequency of alleles in a population; one can only see phenotypes. Nonetheless, the rules can allow one to determine allele frequencies. The following steps allow one to determine the values of p and q in an observed population.

1. Determine the frequency of the recessive phenotype. The only way to have the recessive phenotype is to be homozygous recessive. Thus, the frequency of the recessive phenotype = q^2.
2. Taking the square root of the value determined in step 1 gives the value of q.
3. Since q + p must = 1, the value of p can now be determined.
4. If desired, the values of 2pq and p^2 can also now be determined.

The Hardy-Weinberg equation describes an existing situation. If the five conditions are met, then no change will occur in either allele or genotype frequencies in the population. Of what value is such a rule? It provides a yardstick by which changes in allele frequency, and therefore evolution, can be measured. One can look at a population and ask: Is evolution occurring with respect to a particular gene locus? Since evolution is difficult if not impossible to observe in most natural populations, we will model the evolutionary process using a computer model. The purpose of this simulation is to provide an opportunity to test some of the basic tenets of population genetics and evolutionary biology.

<u>Background on Computer Modeling of Population Dynamics</u>

Evolution occurs in populations of organisms and involves variation in the population, heredity, and differential survival. One way to study evolution is to study how the frequency of alleles in a population changes from generation to generation. In other words, you can ask *What are the inheritance patterns of alleles, not just from two parental organisms, but also in a population?* You can then explore how allele frequencies change in populations and how these changes might predict what will happen to a population in the future.

Mathematical models and computer simulations are tools used to explore the complexity of biological systems that might otherwise be difficult or impossible to study. Several models can be applied to questions about evolution. In this investigation, you will use a program that models how a hypothetical gene pool changes from one generation to the next. This model will let you explore parameters that affect allele frequencies, such as selection, mutation, and migration.

In the second part of the investigation, you will generate your own questions regarding the evolution of allele frequencies in a population. Then you will be asked to explore possible answers to those questions by applying a computer model.

This investigation also provides an opportunity for you to review concepts you might have studied previously, including natural selection as the major mechanism of evolution; the relationship among genotype, phenotype, and natural selection; and the fundamentals of classic Mendelian genetics.

Before coming to lab

We will be using Allele A1 software, developed by Jon C. Herron. The software is free and available to the public. It will be pre-loaded on the computers in the lab, but if you want to download it to your own laptop and use that during class, you may do so. The software is available at: http://faculty.washington.edu/herronjc/SoftwareFolder/AlleleA1.html

You should also watch a YouTube video that will help familiarize you with the software. The video is available at: http://www.youtube.com/watch?v=qGSz4B3s3sg . The video starts out rather out of focus, but that is corrected fairly quickly.

Procedure

1. The default settings in Allele A1 demonstrate Hardy-Weinberg equilibrium. These settings include the assumptions of no mutation, no selection, no migration, no genetic drift (infinite population size), and random mating. In addition, it has initial frequencies of 0.5 for alleles A1 and A2, although those numbers are not necessary for Hardy-Weinberg equilibrium. Run this simulation and convince yourself that these conditions result in no change in the allele frequencies and also that the genotype frequencies can be predicted by the allele frequencies.

   Now, try different values for the starting frequency of allele A1 and run several simulations. Does your experimentation verify that any starting frequencies are in Hardy-Weinberg equilibrium? Explain how you know. (4 points)

2. Now change the population size to 1000 and re-run this simulation (starting with allele A1 and A2 frequencies of 0.5 each). Run this simulation five times. Does the same thing happen each time? Why or why not? You will probably want to select the "Auto" button, which will allow you to compare the different simulations. (4 points)

3. Experiment by changing the initial population size (use 10, 100, and 10,000) and running each of these simulations several times. It may help if you use a different color for each population size. Is the impact of genetic drift similar or different for the population sizes you have modeled? Given what you know about genetic drift, explain your results. (5 points)

4. Under drift, if we track allele A1 for long enough, eventually it will either be fixed (frequency=1) or be lost (frequency=0). You can use this program to quantify how the initial frequency of an allele relates to the probability of either its fixation or its loss. Think about how you can use AlleleA1 to determine this.

   Reset the program to the default values and then change the population size to 100. Run this simulation and record whether allele A1 is either fixed or lost from the population. Ignore those runs in which allele A1 neither becomes fixed nor is lost. Repeat this simulation a total of 30 times, recording each time whether allele A1 is fixed or lost. What % of simulations fixed allele A1? What % lost allele A1? (4 points)

   Now chose a different initial frequency for allele A1 (one that will help you answer the question posed above about how initial frequency relates to its probability of becoming fixed or lost). Repeat this simulation 30 times, again recording the number of times allele A1 is either fixed or lost. What initial frequency for allele A1 did you choose? Given this frequency, what % of simulations fixed allele A? What % lost allele A1? (4 points)

   Finally, chose a third starting frequency for allele A1 to drive home the point. Again, repeat this simulation 30 times, recording the number of times allele A1 is either fixed or lost. What initial frequency for allele A1 did you choose? Given this frequency, what % of simulations fixed allele A1? What % lost allele A1? (4 points)

Now, draw your conclusion to the original question. How does the initial frequency of an allele relate to the probability of either its fixation or its loss? How can you explain that? (6 points)

Individual Investigations

You will now use the Allele A1 software to investigate questions of interest to you. You should plan to do two full investigations in which you investigate a question by altering the settings in the simulation. A full investigation may include running the simulation multiple times, sometimes with the same settings to see if the results are always the same, and sometimes with different settings to see the effect of a given variable. Your questions should be relatively sophisticated. Leaving all the settings at the default values and just changing the initial frequency of A_1 is not a sophisticated investigation. Your second investigation can build on what you found in your first investigation.

Not sure what you want to investigate? Here are some possibilities, but do not feel that you are limited by this list. You can also combine some of these questions to get at specific situations.

- **What is the effect of migration?** Allele A1 is set up to model an island population. The parameter called *Fraction of migrants each generation* determines the number of individuals that move from the mainland to the island every generation, as a fraction of the island population. For example, setting the parameter to 0.1 means that each generation 10% of the individuals in the island population are new immigrants from the mainland. The value you put in here can be quite low (ex: 0.0001). The parameter called *Frequency of A_1 in the source pop'n* determines the frequency of allele A_1 on the mainland (and thus among each generations migrants). Try several different combinations and see what conclusions you can reach.

- **What is the effect of selection?** You can have the fitness of the different genotypes (A_1A_1, A_1A_2, and A_2A_2) vary. Perhaps heterozygotes are the most fit, for example. Perhaps the genotypes that is the most beneficial on the island is not the genotype that is the most beneficial on the mainland, and there is migration. Can migration overcome selection? Your values for fitness can be quite different for the three genotypes (ex: 1.0, 0.4, 0.01) or rather similar (1.0, 0.95. 0.9)

- **What is the effect of population size?** You can run the same parameters multiple times to see if you always get the same result.

- **What happens if an allele is rare?** You can set the starting frequency of A1 to 0.005, for example. Does population size impact the result? Does it matter if the rare allele is beneficial? How beneficial?

- **What happens when the mutation rate changes?** The original settings have the mutation rates equal to zero. What if the mutation rate is 0.00001 in both directions? Is the result consistent? Does population size matter? What if one of the alleles is much more beneficial than the other? What if radiation causes the mutation rate to be closer to 0.01? What if the mutation rate in one direction is different than the mutation rate in the other direction?

- **What changes when you change how many generations you look at?**

Procedure

In your lab notebook, make careful notes about what questions you are investigating, what parameters you used, and what you find in your investigations. Your findings may include final frequencies, or how variable the results were from trial to trial, or how many generations it took before an allele became fixed, or at what point one parameter overwhelmed the effect of another parameter, etc. It really depends on what you decided to investigate. You should write down conclusions that you reach based on the results of the simulations. Refer to actual data when you formulate your conclusions. Depending on your investigation, you may want to graph your results to communicate your findings clearly. Make sure that what you put in your lab notebook will be clear to the reader. Record what you did, what your results were, what you conclude, and why you think you got the results you got. You will be turning in the pages from your lab notebook as well as this packet.