LAB 12 – Natural Selection

Objectives

1. Model evolution by natural selection.
2. Determine allele frequencies within a population.
3. Use the Hardy-Weinberg equation to calculate probability of each genotype in a population.

INTRODUCTION

Gene Pools & Allele Frequencies

When a population of a species evolves by **natural selection**, the overall physical and genetic characteristics of the population change over time due to the collective influence of **selective factors**. Selective factors include anything that impacts survival and reproduction that is not of a random nature, and thus influences the genetic alleles that are passed on to the next generation. Important selective factors in nature include:

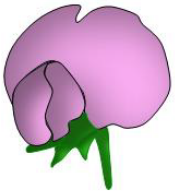
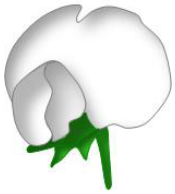
- **climate** (annual temperature and moisture patterns)
- **access to food**
- **pathogens** (disease-causing microorganisms)
- **predation** (predators killing and eating prey)

Over time it is inevitable that selective factors will change, and thus populations affected by these factors will change, i.e. *evolve*, in response.

The physical characteristics of individual organisms are based on genetic alleles, thus the physical characteristics (**phenotypes**) present in a population are based on the genetic alleles in the population (**genotypes**). Evolution of a population by natural selection therefore results in changes in the **gene pool** of the population.

The gene pool includes all alleles in the population for all genes, however we can focus on a single gene to illustrate how the gene pool changes over time due to natural selection. To analyze changes in a gene pool, it is necessary to determine the **allele frequency** of each allele. The frequency of a particular allele is a numerical value representing its proportion among all the alleles for that gene in the population. For example, if in a given population 100% of the alleles for a gene are the **A** allele, its frequency is **1.0**. If 50% of the alleles are **A** and 50% are **a**, then the allele frequencies for each are **0.5**.

To further illustrate how allele frequencies are determined, let's imagine a plant population in which a single gene determines flower color. There are only 2 alleles for this gene in the population, one producing red flower color (**R**) and the other white flower color (**r**), and they exhibit **incomplete dominance**. Thus the phenotypes and corresponding genotypes are:

Phenotype	 red flower	 pink flower	 white flower
Genotype	RR	Rr	rr

If the population contains 65% red-flowered plants, 30% pink-flowered plants and 5% white-flowered plants, the frequencies of the **R** and **r** alleles in the population can be calculated based on the numbers of each allele in a population of 100 individuals (any population size you choose will give you the same allele frequencies, so choose a convenient population size such as 100):

	R alleles	**r alleles**	
65 red (**RR**)	130	0	**R** frequency = 160/200 = **0.80**
30 pink (**Rr**)	30	30	**r** frequency = 40/200 = **0.20**
5 white (**rr**)	0	10	
TOTAL	**160**	**40**	

Notice that each individual contributes 2 alleles to the gene pool, so a population of 100 individuals contains a total of 200 alleles for each gene. Determining the number of **R** alleles in this population and dividing by the *total* number of alleles gives a frequency of **0.80**. Doing the same for the **r** allele gives a frequency of **0.20**. If the calculations are correct, the sum of *all* allele frequencies for a particular gene should equal **1.0**.

In the lab exercise described below, you will keep track of changes in allele frequency in a prey species population over several generations in response to predation. What you should observe is that certain prey phenotypes are more likely to survive and reproduce than others, and thus the genetic alleles responsible for the survivor phenotypes will be passed on to subsequent generations and increase in frequency. Conversely, the alleles carried by prey phenotypes more likely to be eaten should decrease in frequency.

<u>Exercise 1</u> – Predator/Prey simulation

In this exercise, you will recreate the process of natural selection in a simple ecosystem consisting of a single prey species and several predators. The prey species will consist of dried legumes or pasta with 3 different color phenotypes (dark, medium, light). There are 2 alleles for the gene that determines prey color and they exhibit **incomplete dominance** (heterozygotes have a phenotype in between the homozygous dominant and homozygous recessive phenotypes), thus the phenotype reveals the genotype. This will allow you to keep track of changes in the allele frequencies of your prey population in addition to phenotype. Below are the phenotypes and genotypes of all types of prey in your simulation:

<u>color</u>	<u>genotype</u>
dark	BB
medium	Bb
light	bb

You will carry out 3 generations of predation and reproduction (of the prey). In each generation there will be a brief period of predation followed by reproduction (random mating) of the surviving prey. Accounting for all the survivors and their offspring, you will then determine the frequencies of the **B** and **b** alleles. After 3 generations you will determine if the gene pool of your prey population changed, and hence, evolved.

For convenience, we will assume that all matings result in 4 offspring matching the expected phenotypes and genotypes determined by a Punnett square. For example, a **dark (BB) x medium (Bb)** mating will result in **2 dark (BB) & 2 medium (Bb) offspring**. Before you begin the exercise, use the Punnett squares on the first page of your worksheet to determine the 4 offspring produced from each of the 6 possible types of matings.

When you're ready to begin the exercise, designate one member of your group to be a facilitator or "referee" who will coordinate the exercise and enforce the ground rules. The remaining members of your group will be predators attempting to capture prey using a single tool of their choice (spatula, spoon, tweezers, etc.).

Each predator will have a cup into which prey can be "swallowed", and is limited to ***one tool*** and capturing ***one prey individual at a time*** (no fair using hands!). As you proceed, be sure to follow the instructions below, using your worksheet to keep track of all phenotypes and genotypes:

Generation 1:

1. Create a "habitat" for your prey species using the large tray and colored matting provided at your table.

2. The initial population should consist of **25 dark prey** (**BB**), **50 medium prey** (**Bb**), and **25 light prey** (**bb**) scattered within your habitat.

3. Before beginning, calculate the initial allele frequencies for **B** & **b** on your worksheet.

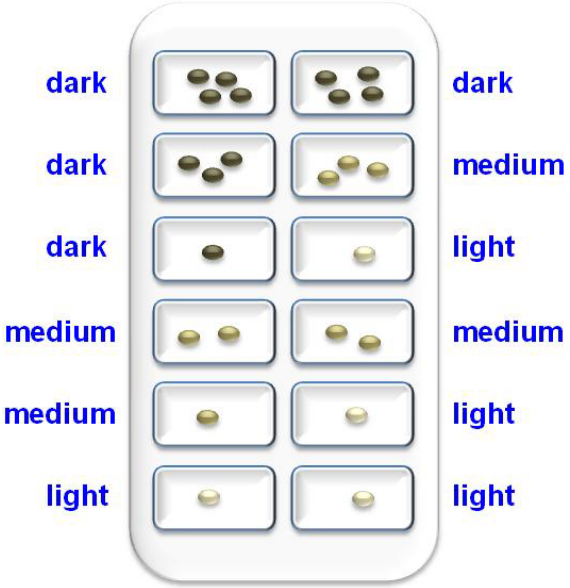
4. The predators will then capture as many prey as they can, **one at a time**, until ~1/3 (~30-35) of the prey remain at which point the facilitator will stop the "feeding".

   NOTE: It is important that ~1/3 of the prey survive in order to maintain your population size at ~100 after the survivors reproduce.

5. Place all the "eaten" prey into the storage containers. Count and record the numbers of surviving prey on your worksheet, then pool them into a cup.

6. Select random pairs of survivors to mate and determine the resulting offspring on your worksheet (unpaired prey go on to the next generation without mating):

   - sort the mating pairs using your ice cube tray as shown below:



Place each member of a mating pair in adjacent slots according to color. For example, for a "dark x light" mating place the dark one on the left of the third row and the light one on the right. In this way you can easily count the number of each type of mating to enter on your worksheet, which in this example would be:

4	dark x dark
3	dark x medium
1	dark x light
2	medium x medium
1	medium x light
1	light x light

   - assume each mating pair produces 4 offspring as determined by Punnett squares

7. Combine the new offspring (as calculated on your worksheet) with the survivors.

8. Accounting for all survivors and offspring, determine the new allele frequencies.

Generation 2:

1. Scatter all remaining prey (survivors & offspring from Generation 1) into the habitat.
2. Predators will collect prey for a period of time determined previously. All prey collected are considered eaten and placed back into the storage containers.
3. Count the surviving prey, record on your worksheet, and pool in a cup.
4. Select random pairs for mating and determine all the resulting offspring on your worksheet (unpaired prey go on to the next generation without mating).
5. Obtain new offspring from the original containers and combine with the survivors.
6. Accounting for all survivors and offspring, determine the new allele frequencies.

Generation 3:

1. Scatter all remaining prey (survivors and offspring) into the habitat.
2. Predators will collect prey for a period of time determined by your facilitator. All prey collected are considered eaten and placed back into the storage containers.
3. Collect and count surviving prey, record on your worksheet, and pool in a cup.
4. Select random pairs for mating and determine all the resulting offspring on your worksheet (unpaired prey will not mate).
5. Accounting for all survivors and offspring, determine the new allele frequencies.

Exercise 2 *(option 1 only)* – Effects of environmental change

For this exercise, continue your predator/prey simulation for 2 more generations *after* changing your habitat to an entirely different color to simulate change in the environment:

Generation 4:

1. Scatter all remaining prey (survivors & offspring from Generation 3) into the habitat.
2. Predators will collect prey for a period of time determined by your facilitator. All prey collected are considered eaten and placed back into the storage containers.
3. Collect and count the survivors, record on your worksheet, and place in a cup.
4. Select random pairs for mating and determine all the resulting offspring on your worksheet (unpaired prey go on to the next generation without mating).
5. Accounting for all survivors and offspring, determine the new allele frequencies.

Generation 5:

1. Scatter all remaining prey (survivors and offspring) into the habitat.
2. Predators will collect prey for a period of time determined by your facilitator. All prey collected are considered eaten and placed back into the storage containers.
3. Collect and count the survivors, record on your worksheet, and place in a cup.
4. Select random pairs for mating and determine all the resulting offspring on your worksheet (unpaired prey do not mate).
5. Accounting for all survivors and offspring, determine the new allele frequencies.

Other Effects on Gene Pools

Genetic Drift

Gene pools will also change randomly due to events that have nothing to do with genuine selective factors, a phenomenon called **genetic drift**. While this includes losing individuals due to random acts of nature such as a being struck by lightning or a volcanic eruption, genetic drift is largely due to the fact that there is an element of chance in what gametes unite in any mating (i.e., which sperm fertilizes which egg). Having studied meiosis and genetics, you realize that meiosis produces huge numbers of gametes in very specific proportions, however the union of gametes when two organisms mate is much like flipping coins or rolling dice – you know the probability but you never really know what you will get. You also should realize that this random deviation from probability is inversely proportional to sample size (recall your coin flips and dice rolls in Lab 10). The only way to avoid such deviations is to have an infinitely large sample size, something that is clearly impossible. So whether you're concerned with rolling dice or producing offspring, there will always be some degree of fluctuation in allele frequencies due to random chance – i.e., genetic drift.

Mutation

Gene pools can also change due to the introduction of new genetic alleles to the gene pool. New alleles arise when the DNA sequence of an existing allele is changed in any way, even by just one nucleotide. Any change in a DNA sequence is called a **mutation**, and mutations can occur as a result of several phenomena: exposure to *chemical mutagens* (carcinogens), exposure to *high energy electromagnetic radiation* (gamma rays, x-rays, UV rays), and *errors in DNA replication, recombination or repair*.

We all accumulate mutations in our somatic (non-reproductive) cells throughout our lifetimes, however a new genetic allele resulting from mutation can enter the gene pool ***only*** if 1) it occurs in a gamete, and 2) the gamete is involved in fertilization that produces a **viable** offspring. If the new allele provides some sort of selective advantage in the current environment, then its frequency will likely increase during subsequent generations.

Gene Flow due to Migration

New genetic alleles can also enter the gene pool when individuals (or the gametes they produce, e.g. plant pollen) of the same species **migrate** into or out of a population. If a migrant produces viable offspring, any new alleles this individual carries may spread in the gene pool. Even if migrants carry no novel alleles, they can still alter the gene pool by changing the frequencies of existing alleles. The collective effects on a gene pool of individuals migrating to and from a population are referred to as **gene flow**.

Hardy-Weinberg Equilibrium *(option 2 only)*

In the early 20th century, Godfrey Hardy and Wilhelm Weinberg identified the conditions necessary for the gene pool of a population to remain stable, and thus for the population to ***not*** evolve. This hypothetical state is referred to as **Hardy-Weinberg equilibrium**. For a population to be in Hardy-Weinberg equilibrium, there must be:

- **no natural selection** (no selective factors acting on the gene pool)
- **no mutation** (i.e., no new genetic alleles)
- **no immigration or emigration** (individuals entering or leaving the population)
- **random mating** (all individuals must have equal opportunities to mate)
- **no genetic drift** (fluctuations in the gene pool due to random chance), which requires an **infinitely large population**

Under these conditions, the allele frequencies in a population will remain unchanged since all alleles will be passed to the next generation via gametes at a frequency equal to the previous generation. By now you should be thinking "sounds great, but those conditions are *impossible* to meet". Exactly! It is inevitable that gene pools will change since there will always be selective factors, new alleles will eventually be generated, mating is rarely random, populations cannot be infinitely large, and random fluctuations (genetic drift) can never be eliminated.

In a sense, this is an argument that gene pools *must* change over time, and therefore populations *must* evolve. Even though no such population can ever exist, the concept of Hardy-Weinberg equilibrium is very useful for making estimates about real populations. For example, if we are interested in a gene for which there are only 2 alleles in a population and we *assume* it is in Hardy-Weinberg equilibrium, the following equation applies:

$$p^2 + 2pq + q^2 = 1$$

This is the **Hardy-Weinberg equation** in which **p** represents the frequency of one genetic allele in a population (e.g., the **B** allele in your predator/prey simulations), and **q** represents the frequency of the other allele (e.g., the **b** allele). Under conditions of Hardy-Weinberg equilibrium, these alleles should be distributed independently among the population based on their allele frequencies.

Although a real population is not in Hardy-Weinberg equilibrium, this equation is nevertheless useful for estimating the *likely* distribution of genotypes if the allele frequencies are known. For example, if the frequency of the **B** allele in your prey population is 0.6, then clearly there is a 0.6 probability that any single allele in the population is a **B** allele. Since each individual is diploid, the probability of an individual inheriting two **B** alleles (**BB**) would be 0.6 x 0.6 ($\mathbf{p^2}$), or 0.36. The same logic applies to the other allele **b**, with a frequency of 0.4, which would yield a probability of 0.4 x 0.4 ($\mathbf{q^2}$) or 0.16 homozygous **bb** individuals. Since there are two ways to be heterozygous (**B** from father & **b** from mother, or **b** from father & **B** from mother), the probability of heterozygous individuals is 0.6 x 0.4 x 2 (**2pq**) or 0.48. Since **BB**, **Bb** and **bb** are the only possible genotypes, the probabilities of each genotype should add up to 1, which they do (0.36 + 0.16 + 0.48 = 1).

Another way to illustrate this equation is to use a variation on the Punnett square. We can place each allele in the population on either side of a Punnett square while indicating the allele frequencies. All possible combinations of gametes (genotypes) in the following generation will be determined, and their probabilities will be the product of the corresponding allele frequencies:

	B (0.6)	**b** (0.4)
B (0.6)	**BB** (0.36)	**Bb** (0.24)
b (0.4)	**Bb** (0.24)	**bb** (0.16)

By this method, you can see how the equation $\mathbf{p^2 + 2pq + q^2 = 1}$ is derived. So if you know the allele frequencies in a population you can use the Hardy-Weinberg equation to estimate the proportion of each genotype, which for our example would be 36% **BB**, 48% **Bb** & 16% **bb**.

The Hardy-Weinberg equation can also be useful if one knows the incidence of a particular genetic condition in a population. For example, cystic fibrosis is due to an autosomal recessive allele and it is known that approximately 1 in 2500 people in the U.K. are afflicted. Since 1/2500 is 0.0004, we can plug this into the equation as follows:

$$\mathbf{0.0004 + 2pq + q^2 = 1}$$

Since those afflicted are homozygous recessive (have 2 recessive alleles), we can say that $\mathbf{p^2 = 0.0004}$, with **p** being the frequency of the recessive allele responsible for cystic fibrosis. The square root of 0.0004 is 0.02, which is the frequency of the cystic fibrosis allele:

$$\mathbf{(0.02)^2 + 2pq + q^2 = 1}$$

If **q** represents the frequency of the normal allele, we know that **p + q = 1** when there are only 2 alleles, so the frequency of the normal allele in the U.K. must be 0.98. Thus the equation is now:

$$(0.02)^2 + 2(0.02 \times 0.98) + (0.98)^2 = 1$$

If we represent the normal and recessive alleles as **F** and **f**, respectively, we can calculate the probabilities of all genotypes in the U.K. population with regard to cystic fibrosis:

FF (normal) = 0.98 x 0.98 = 0.9604

Ff (carrier) = 0.98 x 0.02 x 2 = 0.0392

ff (afflicted) = 0.02 x 0.02 = 0.0004

In summary, since we know that 0.04% (1 in 2500) in the U.K. actually have cystic fibrosis, we can estimate that 3.92% are carriers (1 in every 25.5 people), and 96.04% do not carry the mutant allele at all.

---

*Exercise 2 *(option 2 only)* – Application of the Hardy-Weinberg equation

1. On your worksheet, use the Hardy-Weinberg equation and the allele frequencies you determined at the end of Generation 3 in your predator/prey simulation to determine the *estimated* numbers of each phenotype/genotype (i.e., if the population was in Hardy-Weinberg equilibrium). To do so, determine the probability of each genotype using the Hardy-Weinberg equation, then multiply each probability of each genotype by the total number of individuals at the end of Generation 3 to get the estimated number of each genotype and phenotype. For example, if your allele frequencies at the end of Generation 3 are **B** = 0.4 and **b** = 0.6 and your total population size is **120**:

estimated # of **BB** (dark) = $(0.4)^2$ x 120 = 19.2 or **19 dark**
" **Bb** (medium) = 2(0.4)(0.6) x 120 = 57.6 or **58 medium**
" **bb** (light) = $(0.6)^2$ x 120 = 43.2 or **43 light**

Compare your estimated values to the *actual* numbers of each phenotype to see how well the equation predicts the distribution of genotypes in your population.

2. Use the Hardy-Weinberg equation to answer review questions 2 and 3 on your worksheet.

Before you leave, please make sure your table is clean, organized, and contains all supplies listed below so that the next lab will be ready to begin. Thank you!

Supply List

- Long tray and habitat background matching the color of one prey
- Bag with 3 cups of 3 different color beans representing Light, Medium, Dark colored prey
- Bag with 5 cups and 5 predator tools
- Ice cube tray
- Sorting card

Also, please be sure to do the following before you leave:

- sort all 3 bean colors into the 3 cups and put in prey bag
- put predators and 5 cups in predator Bag

LAB 12 – Natural Selection Worksheet

Name _____

Section_____

Ex. 1 – Predator/Prey simulation

Use the Punnett Squares below to determine the 4 offspring from each mating combination:

(dark) (dark)
BB x BB

___dark : ___medium : ___light

(dark) (medium)
BB x Bb

___dark : ___medium : ___light

(dark) (light)
BB x bb

___dark : ___medium : ___light

(medium) (medium)
Bb x Bb

___dark : ___medium : ___light

(medium) (light)
Bb x bb

___dark : ___medium : ___light

(light) (light)
bb x bb

___dark : ___medium : ___light

Use the tables below to determine the B and b allele frequencies after each generation:

Original population:

color (genotype)	amount	# of B alleles	# of b alleles
dark (BB)	25		----------
medium (Bb)	50		
light (bb)	25	----------	
TOTAL	100		

total # of alleles:	
frequency of B:	
frequency of b:	

Generation 1:

color	# of survivors
dark	
medium	
light	

cross	# of matings	# of OFFSPRING dark	medium	light
dark x dark				
dark x medium				
dark x light				
medium x medium				
medium x light				
light x light				
TOTAL				

survivors + offspring	amount	# of B alleles	# of b alleles
dark (BB)			----------
medium (Bb)			
light (bb)		----------	
TOTAL			

total # of alleles:
frequency of B:
frequency of b:

Generation 2:

color	# of survivors
dark	
medium	
light	

cross	# of matings	# of OFFSPRING dark	medium	light
dark x dark				
dark x medium				
dark x light				
medium x medium				
medium x light				
light x light				
TOTAL				

survivors + offspring	amount	# of B alleles	# of b alleles
dark (BB)			----------
medium (Bb)			
light (bb)		----------	
TOTAL			

total # of alleles:
frequency of B:
frequency of b:

Generation 3:

color	# of survivors
dark	
medium	
light	

cross	# of matings	# of OFFSPRING dark	medium	light
dark x dark				
dark x medium				
dark x light				
medium x medium				
medium x light				
light x light				
TOTAL				

survivors + offspring	amount	# of B alleles	# of b alleles
dark (BB)			----------
medium (Bb)			
light (bb)		----------	
TOTAL			

total # of alleles:
frequency of B:
frequency of b:

Generation 4 *(option 1 only)*:

color	# of survivors
dark	
medium	
light	

cross	# of matings	# of OFFSPRING: dark	medium	light
dark x dark				
dark x medium				
dark x light				
medium x medium				
medium x light				
light x light				
TOTAL				

survivors + offspring	amount	# of B alleles	# of b alleles
dark (BB)			----------
medium (Bb)			
light (bb)		----------	
TOTAL			

total # of alleles:

frequency of B:

frequency of b:

Generation 5 *(option 1 only)*:

color	# of survivors
dark	
medium	
light	

cross	# of matings	# of OFFSPRING: dark	medium	light
dark x dark				
dark x medium				
dark x light				
medium x medium				
medium x light				
light x light				
TOTAL				

survivors + offspring	amount	# of B alleles	# of b alleles
dark (BB)			----------
medium (Bb)			
light (bb)		----------	
TOTAL			

total # of alleles:

frequency of B:

frequency of b:

*Plot the **B** and **b** allele frequencies over the 5 generations of this exercise (be sure to label appropriately).*

(<u>This page is for option 1 only!</u>)

*Describe the overall change in the allele frequencies for **<u>B</u>** and **<u>b</u>** in prey population?*

For how many generations did your most vulnerable prey survive?

What predators (i.e., tools) were most successful at hunting?

<u>Summary of Predator/Prey simulation:</u>

To complete this exercise, examine the allele frequency and phenotype data you have collected to conclude if any changes from your starting population are due to natural selection or genetic drift. <u>Be sure explain the basis of your conclusion and identify any selective factors involved</u>.

<u>Review Questions:</u>

1. *Define the term "gene pool".*

2. *What term is used to indicate that a population has undergone changes in its gene pool?*

3. *For each example below, indicate the associated evolutionary mechanism:*

_____ a change in DNA sequence of an inherited gene

_____ a catastrophic event that kills most of a population

_____ movement of individuals from one population to another

_____ a disease kills vulnerable members of a population

Summary of Predator/Prey simulation *(option 2 only)*:

To complete this exercise, examine the phenotype and allele frequency data you have collected to conclude if any changes from your starting population are due to natural selection or genetic drift. Be sure explain the basis of your conclusion and identify any selective factors involved.

Ex. 2 – Application of the Hardy-Weinberg equation *(option 2 only)*

Using the Hardy-Weinberg equation (**$p^2 + 2pq + q^2 = 1$**)*, calculate the predicted frequencies of each genotype/phenotype in your prey population based on the allele frequencies at the end of Generation 3. Multiply these frequencies by the total population size at the end of generation 3 to obtain "predicted" numbers of each phenotype to compare with the "actual" numbers of each you ended up with:*

color	predicted numbers	actual numbers
dark		
medium		
light		

How do the predicted numbers compare with the actual numbers in your final population?

List the conditions required for Hardy-Weinberg equilibrium.

Does your predator/prey simulation meet the criteria listed above for Hardy-Weinberg equilibrium? If not, what conditions are not met?

Most likely, your actual and predicted numbers above were close but did not match. Might this be related to an unmet condition required for Hardy-Weinberg equilibrium? If so, which one.

Questions for review *(option 2 only)*:

1) If flower color in a plant population is due to 2 alleles which show incomplete dominance, what are the frequencies of each allele if the population consists of 20% red-flowered, 50% pink flowered, and 30% white-flowered plants? **(*show your work!*)**

2) If the frequency of the recessive allele responsible for albinism in the human population is 0.008, what proportion of human beings are albino? are carriers? **(*show your work!*)**

(NOTE: *Dividing 1 by a genotype probability reveals the number of individuals among which there should be 1 individual with that genotype: e .g., if the genotype probability = 0.0001, 1 in 1/0.0001 or 1 in 10,000 have that genotype)*

3) The incidence of phenylketonuria, an autosomal recessive genetic illness due to an enzymatic defect in breaking down the amino acid phenylalanine, is approximately 1 in 10,000 worldwide. What fraction of the human population are carriers? **(*show your work!*)**