Population Genetics --
Evolutionary Stasis and the Hardy-Weinberg Principles[1]

Review and Introduction

Mendel presented the first successful theory of the inheritance of biological variation. He viewed hereditary units as free-floating particles whose integrity is maintained in transfer from parent to offspring. The **chromosome theory of inheritance constrained Mendel's view of genes as free-floating particles** by stating that they are located on chromosomes and **it is the chromosomes, not the genes themselves, that follow Mendel's law of segregation**. Of course, the process of recombination among linked genes relieves this constraint for loci **distantly located** on the same chromosome, and unlinked loci assort independently of one another so that sexual reproduction still produces a tremendous amount of individual variation despite the fact that genes are housed in chromosomes.

As great as these advances were, a complete understanding of the phenomenon of variation required studies at the population level of organization. Population genetics is an extension of Mendelian genetics concerned with the variation existing in a community of reproducing individuals, variously called a **deme**, **local population** or **Mendelian community**. **The basic unit studied in population genetics is the gene pool**, defined as the sum total of genes (both loci and alleles) found in the entire study population.

Local gene pools consisting of individuals who are likely to breed with each other are termed **demes**. We will see that much evolutionary change occurs in these groups.

Modern biologists view each individual is a temporary vessel housing a small fraction of the gene pool. In other words, individuals house only a very small part of the population's total genetic variation. Thus, the concept of a gene pool is an abstract pooling of the genes of all individuals in the population -- it does not exist apart from the individuals themselves.

Within a population, if **mating is random**, the population is said to be **panmictic**. This means that the gametes of each male have an equal chance of fusing with the gametes of each female. Due to **panmixis (= random mating)**, the genes are pooled and the offspring of the next generation are produced by randomly selecting from this pool pairs of alleles for each locus. The situation in nature that most obviously mimics panmixis occurs in many marine invertebrates. For example, corals simply release their gametes into the water.

A final note. To keep things simple, the theory of population genetics we are about to discuss assumes we are dealing with diploid species in which genes are unlinked.

In most cases, are (no linkage, panmixis) likely to be good assumptions? Explain.

The One Gene Model of Population Genetics

Each species has a set number of loci that influence the development of particular aspects of the phenotype. A species may have hundreds of loci and each locus may possess several allelic variants. Although **a given individual** in a diploid species **can only carry at most only two different alleles for each locus (e.g., he or she could be heterozygous)**, there may be many different alleles for some loci in the population as a whole. This allows for a large number of genetically different heterozygotes. As you can easily imagine, the total amount of genetic variation in terms of both loci and alleles in a single population is immense and impossible to analyze in its entirety. For this reason population geneticists use **a simple model** to analyze the genetic nature of a population. Like any model, this oversimplifies reality. But also like any good model, it enables one to grasp the basic principles of genetics at the population level. This model is **called the one gene model of population genetics since it assumes that individuals contain only a single locus**; **it further assumes that this locus has only two allelic expressions.**

Thus, according to this model there is only locus that we will call **A**. For the moment let's assume that there are only two alleles for this locus $\mathbf{a_1}$ and $\mathbf{a_2}$.

Allelic or Gene Frequencies

The structure of **a population is described in terms of the relative frequency of its alleles** (its allelic or **gene frequency**). Even with so simple a situation as that described by the one gene model, it is possible to have variable populations. Let:

$\boldsymbol{p}$ = the frequency of **allele $\mathbf{a_1}$** and

$\boldsymbol{q}$ = the frequency of $\mathbf{a_2}$.

The total frequency of alleles at locus A in the population must equal 1.0 and so:

1. $\boldsymbol{p + q = 1.0}$.

Calculate allelic frequencies just as you would percentages except that the total or whole equals 1.0, not 100. The equation is:

2. frequency of some allele $= \frac{\text{total \# of copies of the allele}}{\text{total \# alleles at found in the pop. at this locus}}$

For example suppose that in a population of 10 individuals there were 3 AA and 2 Aa. Now there are a total of 20 alleles in this population (2 in each individual). Six of these are found in the three AA individuals and 2 in the two Aa. Thus:

freq. A allele $= \frac{(2*\#AA+\#Aa)}{2*\text{total\#individuals}} = \frac{((2*3)+2)}{2*10} = 0.4$

Suppose the following: out of 200 individuals, 50 are AA individuals, 60 Aa, and the remainder aa. Thus, there are 200 - 50 - 60 = 90 aa types.
What are the frequencies of the alleles A and a?
Ans.: A = 0.4 and a = 0.6
Notice a short cut here. If you know that there are only two alleles present in the population for the locus understudy, then you need to use eq. #2 to find the frequency of only one of the two alleles. You can then find the other frequency by subtraction. For example -- suppose that the frequency of allele a_1 is 0.3. Since there are only alleles a_1 and a_2 in the population, then the frequency of allele a_2 = 1.0 - freq($a_{1)}$) = 1.0 - 0.3 = 0.7.

Write an equation for a locus with four alleles. Is it possible for there to be four alleles present in a population? How many alleles for a given a locus are present in one individual?

The Conditions for No Evolution – The Hardy-Weinberg Equilibrium

In 1908 two theorists -- a British mathematician named Hardy and a German physician named Weinberg-- independently discovered that **allelic and genotypic frequencies** which characterize a population **will not change from generation to generation in sexually reproducing populations if a number of conditions are met**. Put another way, Hardy and Weinberg described a series of conditions, which if met, would result in no evolution -- *i.e.*, **STASIS**. This discovery is known as the **Hardy-Weinberg Law** after its co-discoverers.

What are the conditions for stasis? A population will remain stable in what we call a **Hardy-Weinberg Equilibrium** only if:

I. THE ALLELES AT THE LOCUS ARE STABLE -- that is, there is **NO NET MUTATION** of an allele from one form to another. A net mutation rate occurs whenever the **"forward" mutation rate (example: A → a) does not equal the "reverse" mutation rate (here, a → A).**

Using what you have learned earlier in the course, is, for example, it likely that forward and reverse rates will ever be equal?

What is the usual net mutation rate (difference in forward and reverse rates) at most loci?
Are all mutations "visible"?
Is "allelic stability" a "pretty good" assumption? Explain.

II. THE ALLELES ARE NEUTRAL with respect to each other. In other words **NO SELECTION** occurs that favors one allele over another.

Neutral alleles are alleles that are passed on to offspring in the same proportion as they exist in adults because they are **neutral with respect to survival and reproduction in that given environment** ("one's as good as the other"). This assumption is violated when one allele has a higher probability of being represented in the next generation because the individual possessing it has a higher probability of surviving to reproduce. The process of **natural selection** *(our next lecture topic)* results in a violation of this assumption because it confers an adaptive advantage on one allele over another.

The reproductive success of one allele compared to another allele is called its fitness (designated by the letter "**W**"). Neutral alleles possess the same fitness value of 1.0, i. e., they are equal in their reproductive success. Should the environment favor one allele over another (or more precisely individuals possessing one allele over another), then the selectively favored allele will have a fitness value of 1.0 and the fitness value of the other will be a value that is somewhat less.

Give some examples of alleles that are neutral with respect to each other . Consider some of the loci we just considered in the genetics unit? On a molecular level, what sorts of changes in a gene would likely result in no change in selective advantage?
Do you think that it is likely that most alleles have no selective differences with respect to each other?

III. THE POPULATION IS EFFECTIVELY CLOSED: more precisely, there is **NO NET MIGRATION** of one allele into or out of the population. Thus, the Hardy-Weinberg law assumes that populations (demes) are isolated so that individuals neither leave (**emigration**) nor enter from another deme (**immigration**). In an open population, there may be a net **gene flow** from one deme to another. If so, it will disrupt the Hardy-Weinberg equilibrium and cause evolution.

Consider an extreme case in which deme #1 contains only individuals homozygous for a_1 and deme #2 contains only individuals homozygous for a_2. In this example $p = 1.0$, $q = 0.0$ in deme #1 and $p = 0.0$, $q = 1.0$ in deme #2. If just one individual from deme #1 migrates to deme #2 and reproduces, then in the next generation of deme #2 p will be greater than 0.0 and q will be less than 1.0, thus upsetting the Hardy-Weinberg equilibrium.

Does a net migration rate affect the evolution of a population, species or both? Explain.

IV. THE POPULATION IS INFINITELY LARGE AND PANMICTIC. The **Hardy-Weinberg law**, like Mendel's law of independent assortment, **is a statistical law that requires an infinitely large population**. Small populations, like small samples, are subject to **chance deviations from the expected**. The smaller the population, the greater is the probability that the parental gene frequency will not be faithfully reproduced in the offspring generation. Any deviation in gene frequency between these generations due to sampling error is called **GENETIC DRIFT**, which as the name implies is ***nondirectional***. This contrasts with **natural selection** wherein the change in gene frequency is ***directional*** due to an **adaptive advantage** of one allele over another.

Consider a population with only two individuals (one male and one female), both heterozygous for the two allelic expressions of locus A. In this population $p = q = 0.5$. Assume further that this pair produces only four offspring. The Hardy-Weinberg law maintains that the offspring generation will also be characterized by $p = q = 0.5$. According to Mendel's law of segregation, **a monohybrid cross should produce three genotypes in the following ratio: one** a_1a_1 **: two** a_1a_2 **: one** a_2a_2. If this expectation is realized, then the offspring population would have $p = q = 0.5$. How much money would you be willing to bet that the four offspring would follow this expected result? Since **gamete fusion in syngamy is random**, it is possible for all four offspring to be homozygous for the same allele, thus eliminating one allele from the population. Should this occur, then the change in gene frequency between generations would be simply due to chance, or **genetic drift**.

In addition to chance gamete fusion in syngamy, chance deviation in allelic frequency between generations which results in genetic drift could also result from accidental death, chance meeting of mates with different genotypes during the breeding season, and chance variation in number of offspring produced by different individuals in the population. Consequently, there are many specific factors which can result in genetic drift, just as there are many specific environmental factors which act as agents of natural selection to exert a selective pressure on alleles.

PANMIXIS is a bit different. As we said at the start of these notes, it means that **mating occurs at random between the different genotypes in accordance to their frequency**. Thus, if a population consists of 0.2 BB, 0.3 Bb and 0.5 bb, then the chance that a BB individual will mate with a type BB is 0.2 (20%), with a Bb it is 0.3 and with bb, 0.5.

The Hardy-Weinberg Law

You are perhaps most familiar with Hardy-Weinberg as a mathematical formula. We **use this formula to predict the frequencies of the various**

genotypes in a population given the allele frequencies IF AND ONLY IF THE HARDY-WEINBERG POSTULATES (SEE ABOVE) ARE NOT VIOLATED. In such a population, the genotypes will continue indefinitely at the same values, for generation after generation. No evolution occurs, we are in stasis, this is the Hardy Weinberg equilibrium.

On the other hand, **IF A REAL POPULATION WITH A CERTAIN SET OF ALLELE FREQUENCIES HAS GENOTYPES THAT DIFFER FROM THAT PREDICTED BY HARDY WEINBERG, THEN WE KNOW THAT THE POPULATION IS NOT AT EQUILIBRIUM AND EVOLUTION IS OCCURRING BECAUSE ONE OR MORE OF THE POSTULATES HAS BEEN VIOLATED**.

Let's now look at this mathematical expression for a locus where there are just two alleles in the population -- a_1 and a_2.
Let us stipulate that:

***p* = freq (allele a_1)**
***q* = freq (allele a_2)**

Let's remind ourselves of the postulates:

- no net mutation of a_1 to a_2 or vice versa
- no net migration of either allele
- no selection that tends to make possessors of a_1 or a_2 do better relatively speaking in reproduction
- and if the population is so very large that no alleles miss getting into the next generation by chance and alleles are likely to run into each other during fertilization exactly according to their frequencies (panmixis)

Now, the chance that two a_1 alleles come together to form is new individual (Genotype a_1a_1) is equal to the chance of finding one a_1 allele in the population times the chance of finding another. Putting this mathematically:

3. frequency of the **homozygous genotype a_1a_1** in the zygotes = $p * p = \boldsymbol{p^2}$

Likewise:

4. frequency of the **homozygous genotype a_2a_2** in the zygotes = $q * q = \boldsymbol{q^2}$

the **heterozygote, a_1a_2** can be formed two different ways:

An **a_1 meeting an a_2 where the frequency is *p* * *q* , and**
an **a_2 meeting an a_1 where the frequency is *q* * *p***

Therefore,

5. overall the chance of getting a **heterozygote** a_1 a_2, is $(pq + qp) = \boldsymbol{2pq}$

To summarize:

p^2 = the frequency of the homozygous genotype a_1a_1
q^2 = the frequency of the homozygous genotype a_2a_2, and
$2pq$ = the frequency of the heterozygous genotype a_1a_2.

Now recall the equation we saw earlier for allele frequencies in a population:

$$p + q = 1.0$$

If we square that equation, it gives us an expression that tells the probability that two a_1 alleles come together (p^2) ; two a_2 alleles come together (q^2), and the chance of making a heterozygote ($2pq$):

6. $$(p + q)^2 = p^2 + 2pq + q^2 = 1.0$$

This is the famous Hardy-Weinberg expression for a one locus with two alleles present in the population. *Remember, you use this equation to predict the frequencies of the three genotypes from the allele frequencies IF THE POPULATION IS AT HARDY-WEINBERG EQUILIBRIUM*.

A VERY IMPORTANT NOTE

The sum of the frequencies all genotypes in the population must **ALWAYS ADD TO 1.0.** In other words, THE SUM OF THE PARTS EQUALS THE WHOLE.

Showing that the frequencies add to 1.0 only shows you that you calculated the genotype frequencies correctly.

Showing that frequencies add to 1.0 DOES NOT TELL YOU WHETHER YOUR POPULATION IS AT H-W EQUILIBRIUM (unlike what many seem to learn in high school).
Again, you know if you are at the H-W EQUILIBRIUM if the MEASURED GENOTYPE FREQUENCIES MATCH THOSE PREDICTED BY THE H-W EQUATION.

The Hardy-Weinberg Equilibrium -- An Example

Assume a population is in Hardy-Weinberg equilibrium with $p = 0.6$. It follows then that q will equal 0.4. The **predicted genotypic frequencies** will then be: $p^2 = 0.36$, $q^2 = 0.16$, and $2pq = 0.48$. Note that the allelic and genotypic frequencies will each equal 1.0 – they are frequencies after all!

Let's look at how the next generation will theoretically be produced in order to better understand the model. When this generation of adults reproduces, they will produce alleles and genotypes in exactly the same frequency as their

generation. So the next generation will look just like theirs. After all, we said they were at the H-W equilibrium.
For the parental population, if we **assume it consists of the following**:

- 36 homozygous for a_1,
- 16 homozygous for a_2, and
- 48 heterozygotes.

If each individual produces only two gametes, then the number of gametes with a_1 would be 72 + 48 = 120, and the number of gametes with a_2 would be 32 + 48 = 80. The frequency of these 200 gametes would be 0.6 (*p*) for a_1, and 0.4 (*q*) for a_2. The resulting genotypic frequencies among the offspring will be as follows.

	MALE GAMETES (frequency)	
FEMALE GAMETES (frequency)	**a_1 (0.6)**	**a_2 (0.4)**
a_1 (0.6)	**a_1a_1** (0.36)	**a_1a_2** (0.24)
a_2 (0.4)	**a_2a_1** (0.24)	**a_2a_2** (0.16)

After reproduction the offspring genotypic frequency will be the same as the parental generation genotypic frequency.

What if we come back several years later and find the following:
Freq(a_1a_1) = 0.25
Freq(a_1a_2) = 0.70
Freq(a_2a_2) = 0.05
Are the allele frequencies still the same as earlier? Is the population still at H-W equilibrium?
ANS: yes, NO

Microevolution

Evolution refers to a genetic change over time, but what kind of a genetic change? It certainly does not refer to **a change in genes** themselves because this type of genetic change **is called a mutation.** Nor does it refer to a change in an individual's genotype because this does not change over the lifetime of the individual (any somatic mutation - one occurring in a non-reproductive cell - would change a cell's genotype but would not be passed on to offspring). **Individual**s do **change** during their ontogeny (life history) but this **is called development,** not evolution. Species also change over time and if such change is a consequence of a change in their genes, then species change constitutes evolution. But species evolution, also called **macroevolution,** takes thousands of years and is studied by paleontologists, not geneticists.

Definitions of Micro- and Macro-evolution

From a genetic standpoint, ***microevolution is a change in allelic frequency in a single population between successive generations***. This is an important definition because it describes what evolution means to a biologist and emphasizes that **the unit of microevolution is the population (deme) or gene pool**. The term "evolutionary significance", which we have used on occasion since the start of the course, should be interpreted within the context of the definition of microevolution given above and so refers to the significance of some phenomenon **in changing allelic frequency** in a population.

Macroevolution is simply a consequence of microevolution occurring over many generations until a single species changes so dramatically in phenotype that it is classified as a new species.

Evolution results whenever the Hardy-Weinberg law is upset by violating its first four assumptions due to (1) mutation, (2) migration, (3) genetic drift, and (4) natural selection. These four factors result in evolution and so are the causes of evolutionary change.

The study of population genetics establishes evolution as essentially a genetic phenomenon described as a change in allelic frequencies in a gene pool between successive generations. The Hardy-Weinberg law is the foundation of population genetics and establishes the population as the unit of evolution. **This law states that populations are genetically stable over time despite the continual production of individual variation by sexual reproduction.** Only when this genetic equilibrium is disturbed will evolution occur, and there are only four known factors which will upset the Hardy-Weinberg equilibrium: mutation, migration (gene flow), genetic drift and natural selection. Population genetics is essentially the study of how these four factors interact to produce microevolution.

The following is a summary of the basic concepts related to the field of population genetics.

- Microevolution is a genetic phenomenon defined as " a change in allelic frequencies in a single population between successive generations."
- The unit of microevolution is the population, deme or gene pool.
- The Hardy-Weinberg law is the foundation of population genetics and establishes gene pool stability over time despite the continuous generation of individual variation (both phenotypic and genotypic) by sexual reproduction.
- According to the Hardy-Weinberg law, allelic and genotypic frequencies will not change from generation to generation provided the following assumptions are met: genes are stable and neutral, the population is infinitely large and closed. Violation of these four assumptions will result in evolution due to mutation, natural selection, genetic drift and migration, respectively.

- The **ultimate cause** of genetic variation is mutation - the only process whereby new alleles can be formed.
- Sexual reproduction by itself will not result in evolution, i.e., a change in gene frequency, but is important in evolution as the major source (**proximate cause**) of individual variation upon which natural selection acts to produce directional change through adaptation and upon which genetic drift acts to produce random (non-directed) change.
- The one gene model used to explain population genetics is a gross oversimplification of reality in that **it ignores gene interaction in producing phenotypes**. It does, however, provide a basis for defining and examining evolution and makes certain predictions regarding the role of natural selection in changing gene frequencies.

Questions:

1. Define evolution and describe the factors responsible for evolution.

2. Discuss the importance of the Hardy-Weinberg law to population genetics and explain its assumptions.

3. Explain the role of sexual reproduction in the process of evolution. What is the ultimate source of genetic variation upon which sexual reproduction acts to produce individual variation?

4. Distinguish between each of the following:
    a. genotypic and allelic frequency
    b. mutation and migration
    c. gene flow and genetic drift
    d. natural selection and genetic drift
    e. p and q
    f. p^2, $2pq$, q^2
    g. unit of evolution and unit of natural selection
    h. proximate and ultimate sources of genetic variation among individuals

5. Identify each of the following:
    a. gene pool
    b. gene frequency
    c. one gene model
    d. evolution
    e. deme
    f. fitness (W)