Mechanisms of Evolution

36-149 The Tree of Life

Christopher R. Genovese
Department of Statistics
132H Baker Hall x8-7836
http://www.stat.cmu.edu/~genovese/

.

Plan

1. Two More Generations
2. The Hardy-Weinberg Model (Review and More)
3. Populations
4. Blue pages

Announcements

1. Pick up your book, if it's available.
2. Get a reading and writing assignment.

Review: Allele and Genotype Frequency

Consider a hypothetical population and a single genetic locus for a trait with two alleles (A and a).

The *genotype frequency* for this population is the proportion of each genotype:

$$D = \text{frequency of AA} = \frac{\text{\# of AAs in population}}{\text{number of organisms in population}}$$

$$H = \text{frequency of Aa} = \frac{\text{\# of Aas in population}}{\text{number of organisms in population}}$$

$$R = \text{frequency of aa} = \frac{\text{\# of aas in population}}{\text{number of organisms in population}}$$

I use D for *D*ominant homozygote, H for *H*eterozygote, and R for *R*ecessive homozygote.

Review: Allele and Genotype Frequency (cont'd)

The *allele frequency* for this population is the proportion of each allele of the given gene:

$$p = \text{frequency of A} = \frac{2\#\text{ of AAs in population} + \#\text{ of Aas in population}}{2\text{number of organisms in population}}$$

$$= D + \frac{1}{2}H$$

$$q = \text{frequency of a} = \frac{2\#\text{ of aas in population} + \#\text{ of Aas in population}}{2\text{number of organisms in population}}$$

$$= R + \frac{1}{2}H$$

Because there are only two alleles, $p + q = 1$.

If there were more than two alleles for this gene, we would need frequencies for each allele.

The Hardy-Weinberg Model

Assume a population of sexually reproducing organisms satisfying the following: Assumptions

1. Mating is purely random.
2. No flow of genes into or out of the population.
3. No mutation.
4. No differential survival across genotypes.
5. The population is large.

Then, allele (aka gene) frequencies will attain an equilibrium.

Why Is Equilibrium Achieved?

We have defined the following:

Genotype Frequencies	**Allele (Gene) Frequencies**
$D =$ Frequency of AA	$p =$ Frequency of A $= D + H/2$
$H =$ Frequency of Aa	$q =$ Frequency of a $= R + H/2$
$R =$ Frequency of aa	

Under the Hardy-Weinberg assumptions, after one generation of random mating, each individual's alleles are randomly and independently assigned A or a with probabilities p and q respectively.

So,

$$D' = \text{Frequency of AA} = p \cdot p = p^2$$
$$H' = \text{Frequency of Aa} = p \cdot q + q \cdot p = 2pq$$
$$R' = \text{Frequency of aa} = q \cdot q = q^2,$$

and

$$p' = D' + H'/2 = p^2 + pq = p(p+q) = p$$
$$q' = R' + H'/2 = q^2 + pq = q(p+q) = q.$$

The gene frequencies haven't changed!

Mating Pair	Mating Probability	Expected Offspring		
		AA	Aa	aa
AA × AA	D^2	D^2		
AA × Aa	$2DH$	DH	DH	
AA × aa	$2DR$		$2DR$	
Aa × Aa	H^2	$H^2/4$	$H^2/2$	$H^2/4$
Aa × aa	$2HR$		HR	HR
aa × aa	R^2			R^2

Genotype frequencies in next generation:

AA: $D' = D^2 + DH + H^2/4 = (D + H/2)^2 = p^2$

Aa: $H' = DH + 2DR + H^2/2 + HR = 2(D + H/2)(R + H/2) = 2pq$

aa: $R' = H^2/4 + HR + R^2 = (H/2 + R)^2 = q^2$

Total frequency is $p^2 + 2pq + q^2 = (p + q)^2 = 1$.

$$p' = D' + H'/2 = p^2 + pq = p(p + q) = p$$
$$q' = R' + H'/2 = q^2 + pq = q(p + q) = q.$$

The gene frequencies haven't changed!

A Realistic Model?

Do the Hardy-Weinberg assumptions hold for real populations of organisms?

When would they fail most seriously?

When might they apply to good approximation?

Mechanisms of Evolution

Each Hardy-Weinberg assumption that fails provides one way for evolution to take place.

1. Small population

   Chance events can have a long-term impact on gene frequencies.
   Genetic Drift

2. Gene flow into or out of the population

   Direct changes in gene frequencies and an increase or decrease in variation within the population.

3. Mutation

   Generation of novel traits and a source of variation in the population.

4. Differential reproductive rates and mating patterns across genotypes

   Natural Selection

Natural Selection

- *Natural Selection* is a mechanism first proposed by Charles Darwin and Alfred Russel Wallace to explain life's diversity and remarkable adaptation.
- Darwin and Wallace were guided by several observations:
  - There is a struggle for existence in nature.

    Organisms typically produce many more offspring than survive. Limitation of resources and competition prevent exponential growth in populations.
  - Offspring tend to resemble their parent.
  - There is substantial variation in traits within natural populations.
  - An organism's traits can confer advantage (or disadvantage) for survival and reproduction.

Putting these together, we get ...

Natural Selection (cont'd)

As many more individuals of each species are born than can possibly survive; and as, consequently, there is a frequently recurring struggle for existence, it follows that any being, if it vary however slightly in any manner profitable to itself, under the complex and sometimes varying conditions of life, will have a better chance of surviving, and thus be naturally selected. From the strong principle of inheritance, any selected variety will tend to propagate its new and modified form.

— Charles Darwin, *On the Origin of Species* (1859)

Natural Selection (cont'd)

At it's heart, natural selection is a logical argument.

Suppose that the following are present in a population under study:

1. Reproduction.
2. Heredity.
3. Variation in heritable characters.
4. Differential reproductive successive across different types.

Then those with greater reproductive success will produce more offspring, who will tend to resemble them in those characters that confer reproductive advantage, and who will thus tend to propogate.

This will produce cross-generational change in the population: evolution.

A Simple Selection Model

- Now, we'll consider the same model with a small tweak.
- Use the Hardy-Weinberg assumptions with the addition that each genotype now survives with the following probabilities:

Genotype	Survival Probability
AA	1
Aa	1
aa	$1 - u$

  where u is a number between 0 and 1.

- What do you predict will happen to the gene (allele) frequencies in a population governed by this model?

Activity (cont'd)

Start with the set of cards you finished with last time.

Perform the same steps again with one exception: assume that homozygous recessive (aa) offspring do not survive.

Set aside any aa pairs you get.

After One Generation . . .

Think of the selection model as producing the next generation by random mating followed by elimination of the aa offspring.

After the random mating, we have (from earlier) that

Eliminating the aa's means reducing q^2 to $(1-u)q^2$.

$$\begin{aligned} D' &= \text{Frequency of AA} = p^2 \\ H' &= \text{Frequency of Aa} = 2pq \\ R' &= \text{Frequency of aa} = q^2. \end{aligned} \longrightarrow \begin{aligned} D' &= \text{Frequency of AA} = p^2 \\ H' &= \text{Frequency of Aa} = 2pq \\ R' &= \text{Frequency of aa} = (1-u)q^2. \end{aligned}$$

$$\swarrow$$

But these are relative frequencies: $D' + H' + R' = 1 - uq^2 < 1$.
Divide by the sum to get true proportions.

$$\begin{aligned} D' &= \text{Frequency of AA} = \frac{p^2}{1-uq^2} \\ H' &= \text{Frequency of Aa} = \frac{2pq}{1-uq^2} \\ R' &= \text{Frequency of aa} = \frac{q^2-uq^2}{1-uq^2}. \end{aligned}$$

After One Generation . . . (cont'd)

Because $p' = D' + H'/2$ and $q' = R' + H'/2$, it follows that

$$p' = \frac{p^2 + pq}{1 - uq^2} = \frac{p}{1 - uq^2}$$

$$q' = \frac{(1-u)q^2 + pq}{1 - uq^2} = \frac{q - uq^2}{1 - uq^2}$$

After each generation, the frequency of A has increased by a factor of $1/(1-uq^2)$, which is greater than 1 if $u > 0$.

Therefore, A will eventually be the only allele remaining.

Fitness

At the each generation in the selection model, the three genotypes have different rates of reproduction.

For example, in the first generation

Genotype	**Mate**	**Probability**	**Expected Offspring**		
			AA	Aa	aa
AA	AA	D	1	0	0
	Aa	H	0.5	0.5	0
	aa	R	0	1	0
Aa	AA	D	0.5	0.5	
	Aa	H	0.25	0.5	0.25
	aa	R	0	0.5	0.5
aa	AA	D	0	1	0
	Aa	H	0	0.5	0.5
	aa	R	0	0	1

On average, only $1 - u$ of the aa offspring survive from each mating.
All of the others survive.

Fitness (cont'd)

Weighting each possible mating by its probability, we get the expected number of offspring for each genotype.

$$\begin{aligned}
\text{offspring(AA)} &= 1 \\
\text{offspring(Aa)} &= D + \frac{3}{4}H + \frac{1}{4}H(1-u) + \frac{1}{2}R + \frac{1}{2}R(1-s) \\
&= D + H + R - \frac{u}{2}(R + \frac{1}{2}H) \\
&= 1 - \frac{u}{2}q \\
\text{offspring(aa)} &= D + \frac{1}{2}H + \frac{1}{2}H(1-u) + R(1-u) \\
&= D + H + R - u(R + \frac{1}{2}H) \\
&= 1 - uq.
\end{aligned}$$

Fitness (cont'd)

Natural selection indicates that those phenotypes/genotypes with greater *reproductive success* will propogate in the population.

The term *fitness* is used as a measure of reproductive success of each genotype relative to the others.

There are different ways to define it, but in this case take the fitness of a genotype as ratio of the expected number of offspring for that genotype over the comparable number for AA.

From the above calculations, we get

$$\begin{aligned} \text{fitness(AA)} &= 1 \\ \text{fitness(Aa)} &= 1 - \frac{u}{2}q \\ \text{fitness(aa)} &= 1 - uq. \end{aligned}$$

Eventually AA will take over the population!

Activity

Setup

1. From the gene pool, select 10 cards, and then go back to your seat.
2. Count how many 'A' alleles you have and how many 'a' alleles you have and record these.
3. Combine the totals with your class mates (two by two, then four by four, etc.) to get a total count for the class.
4. Calculate the frequencies of A and a in the class and record these numbers.
5. Now *randomly* pair your alleles so that you have 5 gene pairs just like the genotypes of diploid organisms.
6. Count how many 'A' alleles you have and how many 'a' alleles you have and record these.
7. Combine the totals with your class mates (two by two, then four by four, etc.) to get a total count for the class.
8. Calculate the frequencies of each of these combinations in the class and record.

Activity

Mating: The First Generation

9. Spread your 10 cards out in one hand as if you were playing cards. Make sure that you are the only one who can see the alleles. Go around the room and find another person to trade with. Let that person take one of your alleles. You take one of their alleles. (Don't peek!) Now take your new allele and put it away so that you don't trade it again.
10. Trade 4 more alleles with 4 more people (only 1 trade with each person).
11. When you have only 5 alleles left in your hand and 5 new ones that are trades, sit down!
12. Again, pair up your 10 alleles randomly so that you have 5 pairs.
13. Count how many of each allele and each genotype you have in your hand;
14. Combine with classmates as bfore, calculate the frequencies of each of these combinations and record.

Mating: Rinse and Repeat. (It never gets old!)