

7 POPULATION GENETICS

7.1 INTRODUCTION

Most humans are susceptible to HIV infection. However, some people seem to be able to avoid infection despite repeated exposure. Some resistance is due to a rare allele of a gene that codes for the CCR5 protein. CCR5 is a cell surface protein that is a co-receptor for the HIV virus when it binds to the cell membrane of macrophages and T cells. Some people have a form of the gene encoding CCR5 that includes a 32-bp deletion that results in a dysfunctional protein. The allele is called CCR5- Δ 32. Individuals who inherit two copies of the deleted allele have no CCR5 co-receptors on the cell membrane of white blood cells. These homozygotes are highly resistant to HIV infection. Will the global HIV epidemic result in increasing frequencies of the CCR5- Δ 32 allele? If so, how soon will this happen? Also, northern European populations carry a higher frequency of this allele than do Asian or African populations. Is this difference due to chance, or due to the fact that CCR5- Δ 32 also increased resistance to some previous epidemic that occurred in Northern Europe (plague, smallpox, etc.)?

These are kinds of questions that **population genetics** is designed to answer. This discipline describes the behavior of alleles in populations by focusing on the forces that can cause allele frequencies to change over time. "Allele frequency change over time" is simply a definition of "evolution". So population genetics is that branch of genetics that is concerned with the evolutionary processes of natural selection, genetic drift, mutation, migration, and non-random mating. Population genetic approaches can be used to understand the consequences of these processes individually or in combination.

Review: A **population** is a group of organisms of the same species living within a prescribed geographical area. The area is usually determined to be of a size within which individuals are likely to find mates. Geographically widespread species are often subdivided into more or less distinct breeding groups that live within limited geographical areas. These groups are called subpopulations. For example, Ponderosa pine in Southeastern Arizona and Northern Mexico could be considered a single population. Individual pines in Arizona live at high altitudes, but not in desert surrounding the mountains. If there is only rare migration (pollen dispersal) between groups in different mountain ranges, each mountain range would have its own local subpopulation of pines. Humans were probably once subdivided into local populations. Now, for the most part, we exist in a global population (with the exception of a few fairly isolated subpopulations).

The complete set of genetic information contained within the individuals in a population is called the **gene pool**. The gene pool includes all genetic loci and all the alleles for each locus present in the population. The gene pool of a population can be described and characterized by certain statistical properties:

Allele frequency (gene frequency) is the proportion of all alleles at a locus that are of a specified type.

Polymorphism is the proportion of all genetic loci that exist in more than one allelic form. Humans, $P=0.32$, *D. melanogaster*, $H=0.42$.

Heterozygosity is the proportion of individuals that are heterozygous (as opposed to homozygous) for alternate alleles at a specific locus. Humans: $H=0.06$, *D. melanogaster*, $H=0.14$

Both heterozygosity and polymorphism are used as measures of genetic variation within populations. Different alleles within populations can be identified by phenotypic effects (color, size, shape variants, disorders like color blindness or Huntington's disease), by variation in protein structure (standard blood-type assays, protein electrophoresis), or by variation in DNA sequence (restriction site variation, variation in number of repeats in repetitive DNA segments, DNA sequence variation)

We will begin our discussion of population genetics by considering an analysis of a local population with respect to a phenotype determined by two alleles at a single locus. The human MN blood group is an extremely simple system. It is characterized by a single locus with only two alleles, M and N (sometimes referred to as L^M and L^N). The alleles are codominant (both alleles are detectable in heterozygotes).

blood types (phenotypes)	Genotypes
M	M M (or $L^M L^M$)
MN	M N
N	N N

ALLELE & GENOTYPE FREQUENCIES IN SURVEY OF BRITISH POPULATION:

Phenotype	Genotype	# M alleles	# N alleles
298 M	298 MM	596	0
489 MN	489 MN	489	489
213 N	213 NN	0	426
Total 1000	1000	1085	915

Genotype frequency is relative proportion of genotype:

Proportion of MM genotype = $298/1000 = 0.298$

Proportion of MN genotype = $489/1000 = 0.489$

Proportion of NN genotype = $213/1000 = 0.213$

Total = 1.000

The allele frequency of the M allele is the relative proportion of the allele:

Proportion of M allele = $1085/2000 = 0.5425$

Proportion of N allele = $915/2000 = 0.4575$

Total = 1.000

If we let p = the allele frequency of M and q = the allele frequency of N, then $q = 1 - p$.

7.2 THE HARDY-WEINBERG PRINCIPLE

We can use knowledge of allele frequencies to make inferences about patterns of mating, selection on certain alleles, migration between populations, etc. Allele frequencies are more useful than genotype frequency because alleles rarely undergo mutation in a single generation, so are stable in their transmission from one generation to the next. In contrast, genotypes are not permanent. They are broken up by the processes of segregation and recombination that take place during meiosis. Furthermore, we can deduce the expected genotype frequencies in the next generation from knowledge of only the allele frequencies in the previous generation.

First, consider that for diploid organism with two different alleles at a locus, a gamete has an equal chance of containing either of the two alleles (equal segregation--Mendel's first law).

individual	gamete types	gamete contribution	gamete frequency
0.298 MM	all M	0.298 M gametes	0.5425 M gametes
0.489 MN	1/2 M 1/2 N	0.2445 M gametes 0.2445 N gametes	
0.213 NN	all N	0.213 N gametes	0.4575 N gametes

Next assume random mating of individuals with respect to MN blood type (usually a very reasonable assumption, as people do not usually chose their spouses based on MN blood type!). Random mating is equivalent to random union of gametes.

One Generation of Random Mating (Punnet Square Approach)

Gametes		Eggs		Allele frequencies next generation
Sperm		0.54 M	0.46 N	
(p=0.54)	0.54 M	$p^2=0.2943$ MM	$pq=0.2482$ MN	$p'=0.2943+0.1241+0.1241=0.5425$
(q=0.46)	0.46 N	$pq=0.2482$ MN	$q^2=0.2093$ NN	$q'=0.2093+0.1241+0.1241=0.4575$

THEREFORE:

1) Zygotic genotype frequencies are predictable from gamete frequencies, assuming random mating, and

Allele frequencies do not change from generation to generation under this scenario
 Genotype frequencies are given by the following rules:

- frequency of MM = p^2
- frequency of MN = $2pq$
- frequency of NN = q^2

2) Also note that $p^2+2pq+q^2=1$.

A common mistake is to believe that if an allele is dominant, it should become more frequent over time. Hardy-Weinberg principle tells us that, unless some outside force is operating to change allele frequencies, they will remain constant from generation to generation.

HARDY-WEINBERG PRINCIPLE:

p^2	$2pq$	q^2
MM	MN	NN

Although the H-W principle is quite simple, it has a number of important implications:

Allele frequencies remain constant from generation to generation (unless some force is acting to change them).

If genotypes have equal viability, the genotype frequencies in adults will be the same as those in the zygotes. The allele frequencies will then be unchanged from those in the first generation.

$$p' = p^2 + 2pq/2 = p^2 + pq = p(p + q) = p$$

Although this is a simple result, we had to make a number of assumption in order to arrive at it:

- Mating is random across the entire population—no subpopulations that differ in allele frequencies.
- Allele frequencies are the same in females and males.
- All genotypes have equal viability and fertility (no selection).
- Mutation does not occur, or is so rare it can be ignored (no mutation).
- Migration into the population can be ignored (no migration)
- The population is large enough that the allele frequencies do not change from generation to generation due to chance (no random genetic drift).

Deviations from these assumptions can lead to changes in allele frequencies from one generation to the next. These assumptions therefore describe the potential forces that cause allele frequencies to change through time. Therefore, if we can measure deviations from H-W equilibrium in populations, we can begin to determine if allele frequencies are changing, and also begin to measure the forces causing this change.

For Example, the British blood group data:

$$\text{freq M} = p = 0.54$$

$$\text{freq N} = q = 0.46$$

Expected genotypic frequencies:

MM	MN	NN
p^2	$2pq$	q^2
0.294	0.496	0.209

So, if 1000 individuals are sampled in the following generation,

Genotypes	MM	MN	NN
Expected	294	496	209
Observed	298	489	213

O	E	$(O-E)^2$	$(O-E)^2/E$
298	294	16	0.054
489	496	49	0.099
213	209	16	0.077
	d.f. = 1		$\chi^2 = 0.230$

Since there is no significant deviation from the H-W expectations, we conclude that assumptions i-vi are justified.

When an allele is rare, there are many more heterozygotes than there are homozygotes ($2pq \gg q^2$ if q is small).

For example, cystic fibrosis is an inherited disorder of the pancreas and lungs. It is one of the most common recessive genetic disorders in Caucasian populations, occurring in 1/1700 newborns. Heterozygotes are not easily identified, but we can calculate how many heterozygous carriers occur by using the H-W principle:

AA	Aa	aa	(affected individuals are aa)
p^2	$2pq$	q^2	

We know that $q^2 = 1/1700 = 0.0006$ (frequency of the aa genotype)

So,

$$q = \sqrt{0.0006} = 0.024 \text{ (frequency of the a allele)}$$

$$p = 1 - q = 0.976 \text{ (frequency of the A allele)}$$

Therefore, the frequency of heterozygous carriers is

$$2pq = 2 * 0.024 * 0.976 = 0.047.$$

This implies that 1 out of every 21 people is a heterozygous carrier of the gene for cystic fibrosis.

That is, mating is random wrt to MN blood type, individuals of different genotypes are equally viable and fertile, there is no substantial population substructuring wrt these alleles, mutation is rare, and the population is large enough that there is no substantial effect of genetic drift.

3. The same principals can be applied to sex-linked traits if males and females are treated separately. Males have only one copy of a sex-linked gene so for males: genotype frequencies = allele frequency.

Genotype	Allele frequency	Genotype frequency
$X^A Y$	p	p
$X^a Y$	q	q

Consider a recessive sex-linked gene causing color blindness. If 8% of men are color blind, then the allele frequency can be directly calculated as $q=0.08$. Then the H-W principle can be applied to calculate the expected frequency in women:

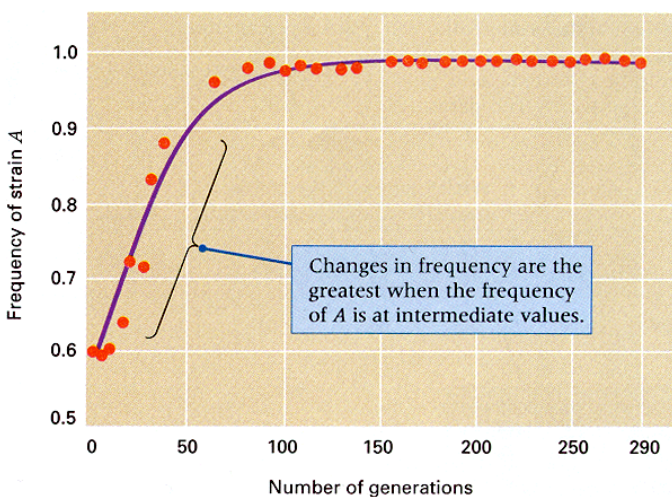
Genotype	Expected Frequencies	
$X^A X^A$	p^2	0.8464
$X^A X^a$	$2pq$	0.1472
$X^a X^a$	q^2	<u>0.0064</u>
Total		1.0000

So while 8 out of 100 men are color blind, fewer than 1 out of 100 women are color blind.

7.3 NATURAL SELECTION

In each generation, genotypes that promote survival or reproduction in the prevailing environment are more common among individuals that are reproducing. These genotypes therefore contribute disproportionately to the offspring of the next generation. Alleles that increase survival or reproduction increase in frequency from generation to generation (like the allele conferring tetracycline resistance in the example in the previous section). Over multiple generations, the individuals within a population become better able to survive and reproduce in the prevailing environment. This is the process of adaptation by natural selection. We will consider two categories of selection in this class: directional selection and balancing selection. In population genetics, **directional selection** refers to cases in which one allele (or one homozygous genotype) always has the highest fitness. **Balancing selection** refers to cases in which either 1) a combination of different alleles (e.g., a heterozygote) has the highest relative fitness, or 2) when the fitnesses of different alleles or genotypes depends on environmental conditions or on the genotypic composition of the population.

7.3.1 DIRECTIONAL SELECTION IN A HAPLOID ORGANISM. SELECTION IN E. COLI IN A LACTOSE-LIMITED ENVIRONMENT

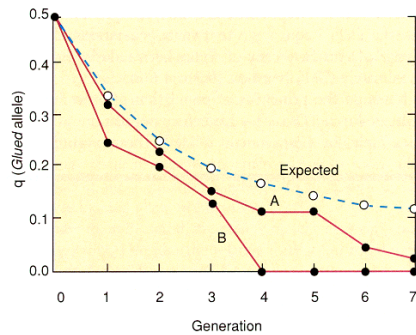


This figure shows the results of competition between two bacterial genotypes, A and B. Genotype A has a constitutive mutation for enzyme production in the lactose operon. B is the normal inducible lactose operon. Both strains were grown together in an environment with no glucose and limited lactose. Genotype A gradually replaced genotype B due to its better ability to compete for limited lactose. In this experiment, the competition was allowed to continue for 290 generations, during which time the proportion of the A genotype (p) increased from 0.6 to 0.9995, and that of B

decreased from 0.4 to 0.0005. Changes in allele frequency over time is one definition of evolution. This bacterial population has therefore undergone evolution from one allelic state (A and B about equally common) to another (B very, very rare) in the course of 290 generations (about 6 days).

Experiments such as this allow us to observe the evolutionary process, and to study evolutionary forces in fine detail. Most of the theoretical and empirical advances in population genetics have been made by biologists interested in understanding the evolutionary process.

7.3.2 DIRECTIONAL SELECTION IN DIPLOID ORGANISMS.



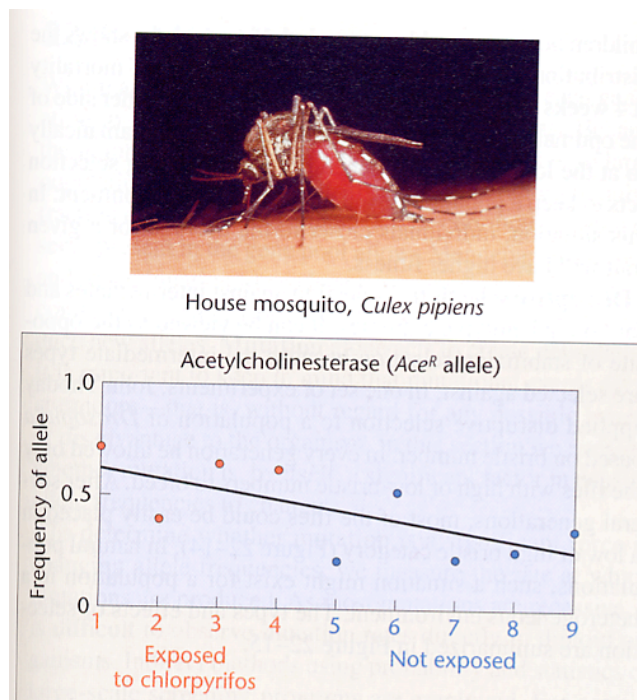
A) *Glued* is a wing mutation that is deleterious when homozygous.

In two replicates (A and B) the frequency of the *Glued* allele at the beginning of the experiment was 0.5.

After 7 generations, the allele frequency had decreased to less than 0.025. The dashed line shows the expected rate of elimination of a lethal gene in a diploid population. The evolution of this laboratory population from one allelic state to another took about 3 months.

FIGURE 19.15 The expected change in the frequency of the lethal allele *Glued* (blue line) and the observed change in the replicate populations (red lines).
Source: Data from Clegg, et al., *Genetics*, 83:793–810, 1976.

B) Exposure to insecticide changed the allele frequencies at the acetylcholinesterase gene in the house mosquito.



7.3.3 *DIRECTIONAL SELECTION ON RECESSIVE AND DOMINANT ALLELES IN DIPLOIDS*

Genotype	AA	Aa	aa
Genotype Frequencies	p^2	$2pq$	q^2
Genotypic Fitness:	W_{AA}	W_{Aa}	W_{aa}
Relative Fitness	W_{AA}/\bar{W}	W_{Aa}/\bar{W}	W_{aa}/\bar{W}

$$(\bar{W} = \text{average fitness} = p^2W_{AA} + 2pqW_{Aa} + q^2W_{aa})$$

A. Freq of a after one gen. of selection: $q' = \text{freq}(aa) + 1/2 \text{freq}(Aa) = q^2 W_{aa}/\bar{W} + pq W_{Aa}/\bar{W}$

B. Freq of A after one gen. of selection: $p' = \text{freq}(AA) + 1/2 \text{freq}(Aa) = p^2 W_{AA}/\bar{W} + pq W_{Aa}/\bar{W}$

So one can predict the allele frequency in a future generation if the current allele frequency and the fitness values are known.

Selection on rare recessives is not very efficient:

Selection for or against rare recessive alleles is inefficient because:

- a) Almost all individuals with the allele are heterozygous.
- b) Heterozygotes for the recessive allele have same phenotype (and same fitness) as homozygotes for the dominant allele.

That is, if a is a **rare recessive** allele, then p will be ≈ 1 , and q^2 will be a very small number, meaning that there will be very few individuals of genotype aa in the population. Also $W_{AA}=W_{Aa}$. Since there are no aa individuals, the mean fitness of the population will be close to the fitness of the dominant phenotype ($\bar{W} \approx W_{AA}=W_{Aa}$).

The change in frequency of the a allele due to selection will then be:

$$q'-q \approx [pqW_{Aa}/\bar{W}] - q \approx [pq] - q \approx q - q \approx 0$$

Therefore selection is not effective because it does not cause allele frequency to change much (at least in the short term).

Selection on rare dominant alleles is efficient:

But selection will be effective for a dominant allele. The change in frequency is determined by the ratio of the fitness of the A phenotype to the average fitness of the population.

If A is a **rare dominant** allele, then $q \approx 1$ and p^2 is very small. The mean fitness of the population will be close to W_{aa} ($\bar{W} \approx W_{aa}$).

The change in frequency of the A allele will be:

$$p'-p \approx [pqW_{Aa}/W_{aa}] - p \approx [pW_{Aa}/\bar{W}] - p$$

Q: what will happen if the A allele is advantageous, so that $W_{Aa} > \bar{W}$?

Q: Will the *CCR5* $\Delta 32$ allele increase in human populations? See discussion in Freeman and Herron (p. 163-165).

7.3.4 THE EVOLUTION OF DRUG RESISTANCE IN PATHOGENS

We have already discussed the evolution of antibiotic resistance in tuberculosis and other bacterial populations. In addition, the evolution of resistance to antimalarial drugs in the protozoan that causes malaria is becoming an increasing problem in places where malaria is endemic.

7.3.5 THE EVOLUTION OF PESTICIDE RESISTANCE

In addition to the examples of the evolution of drug resistance in pathogens, there are other examples of the practical importance of understanding the principles of population and evolutionary genetics. The last 50 years have shown the rapid development of pesticide resistance in many pest species:

yellow fever mosquito
Colorado potato beetle
Norway rat

Many conditions of modern society contribute to pest problems:

1. agricultural monocultures (easy dispersal of pests and pathogens, little genetic variation for pest or disease resistance)
2. concentration of human populations in cities

After WWII, synthetic organic pesticides were introduced to control agricultural and public health pests. These pesticides had the immediate effect of greatly increasing crop production and decreasing the incidence of many crop diseases. However, many pests rapidly evolved resistance to pesticides. The following is a partial list of cases in which pests that were originally controlled by organic pesticides have now developed resistance to these substances.

1. In 1969 (after about 20 years of DDT use) 15 different species of malaria mosquitoes had evolved resistance to DDT. In 1976, 43 species were resistant.
2. In India 1/2 million cases of malaria in 1969, but 27 million cases in 1977
2. Houseflies, flour beetles, seven species of rodents have also developed resistance to various pesticides
3. Since 1970, about 50 different species of weeds have evolved resistance to herbicides.
4. Warfarin resistance in rats evolved after about 15 years of the use of the anticoagulant Warfarin

7.3.6 BALANCING SELECTION

In a locus with two alleles, we have so far assumed that one of the homozygotes has higher fitness than the other, and that the heterozygote has the same fitness as one of the two homozygotes. It is also possible that the heterozygote has fitness intermediate between the two homozygotes. In either of these cases, the allele associated with the superior homozygote will increase in frequency and will become fixed (frequency=1) unless it is opposed by mutation pressure (in which case it will have a frequency close to, but slightly less than 1).

However, if the heterozygote fitness is greater than the fitness of either homozygote, both alleles will be maintained at high frequency by selection. In each generation, the heterozygotes produce more offspring than the homozygotes, and selection for heterozygotes keeps both alleles in the population.

This phenomenon, called **heterozygote advantage** or **overdominance**, does not appear to be very common in natural populations. The best-known example is that of the sickle-cell anemia mutation (Hb^S), and its relation to a type of malaria caused by a parasitic protozoan (*Plasmodium falciparum*).

Without medical care, homozygotes for Hb^S usually die while quite young. Yet in parts of Africa and the Middle East, the frequency of this allele reaches 10% or higher. The allele is maintained at high frequency in these populations because heterozygotes are less susceptible to malaria, and have milder infections, than people who are homozygous for the wild-type allele. Heterozygotes therefore have the highest fitness in these populations.

Because there are so few examples of alleles demonstrating heterozygote superiority, it is not thought to contribute greatly to the genetic variation that is observed within population. However, there are other types of balancing selection that may contribute a large fraction of the standing genetic variation in populations. **Frequency-dependent selection** (where any allele that is rare gains a selective advantage simply because of its rarity), **genotype-environment interaction for fitness**, in which the fitness of a genotype changes depending on the environment, and **antagonistic pleiotropy** in which alleles can have deleterious effects on one aspect of fitness, but beneficial effects on another aspect of fitness (e.g. tradeoffs between fertility and survival) all can maintain genetic variation within populations.

Because of possible consequence of loss of genetic variation from small population sizes, selective breeding in crops and livestock, or to inbreeding, it is important to determine whether most variation is maintained by selective processes like these. If it is, then the consequences of loss of genetic variation could be disastrous for a population (or an agricultural species).

Q: If most variation is not maintained by these selective forces, what does maintain it?

7.4 MUTATION

We can measure the rate of mutation for certain kinds of mutations, to determine whether this is a powerful force for changing allele frequencies between one generation and the next. Achondroplasia (a kind of dwarfism) is determined by a dominant allele (D), so the occurrence of the condition in the offspring of two normal parents provides an estimate of the mutation rate. In about 3 births out of every 100,000 a mutation to the achondroplasia allele occurs. This implies a mutation rate of $3/200,000 = 1.5 \times 10^{-5}$. (200,000 gametes contributed)

In a population in which the initial frequency of the D allele is 0.0, $q=0$ and $p=1$. The frequency of D after one generation of mutation is

$$q' = p \cdot 1.5 \times 10^{-5} = 1.5 \times 10^{-5} = 0.000015$$

The frequency of the alternate allele (d) will therefore be

$$p' = 1 - 1.5 \times 10^{-5} = 0.999985$$

So the effect of mutation on allele frequencies over a few generations is very small.

However, mutation is very important over the long term. Mutation is the ultimate source of all genetic variation. Without mutation, there would be no physical basis for the differences seen between individuals in a single population, or for the differences seen between species. Although the overwhelming majority of mutations that occur are thought to be deleterious, advantageous ones do sometimes occur. Consider, for example, a mutation that allows a *Streptococcus pneumoniae* bacterium to live and reproduce in the presence of a common antibiotic like tetracycline. In an environment where tetracycline is common (a hospital for example) tet^R bacteria might have a competitive advantage over tet^S bacteria. Over the course of several generations, the tet^R bacteria would then become common, where it once had been very rare. The bacterial population is then said to have evolved tetracycline resistance.

The evolution of antibiotic resistance has recently become a very serious public-health concern. In many major urban hospitals in the United States, antibiotic-resistant tuberculosis is becoming very common.

Malaria parasites are also becoming resistant to many of the anti-malarial drugs that have previously been very effective in treating this disease.

7.4.1 EFFECT OF RECURRENT MUTATION.

Even though selection on rare recessives is inefficient, it will tend to decrease the frequency of rare deleterious recessives over a large number of generations. For example, a recessive lethal at an initial frequency of 0.02 will be reduced in frequency to 0.01 in 50 generations, if there is no recurrent mutation. However, if there is recurrent mutation to the deleterious recessive allele, there will be an equilibrium frequency of the allele, where the increase in frequency due to recurrent mutation is exactly matched by the decrease in frequency due to selection. For a recessive allele, the equilibrium allele frequency is:

$$q_e = \sqrt{(u/s)}$$

where u is the mutation rate and s is the selection coefficient against the mutation (s is the amount by which the fitness of a homozygous recessive individual is reduced below that of an individual with the dominant phenotype). Since the mutation rate is generally much smaller than the selection coefficient, q_e is usually a small number. For example, a recessive lethal mutation has a selection coefficient of 1. If the mutation rate is 10^{-5} , then $q_e = \sqrt{(u/s)} = \sqrt{(1/10^{-5})} = 0.0001 = 0.003$. So, even very strongly disadvantageous mutations like lethals can have a non-negligible frequency in populations due to recurrent mutation.

Much of the genetic variation present in populations is thought to result from just such a balance between selection and mutation. One hotly-debated topic in population genetics in recent years has been the question of whether mutation accounts for most of the variation present within populations, or whether some variation is maintained by other forces. Although we have so far only considered selection that removes variation from populations, some types of selection can actually maintain variation.

7.5 NONRANDOM MATING (INBREEDING)

7.5.1 INBREEDING REDUCES HETEROZYGOSITY

The most well-known form of non-random mating is inbreeding. Inbreeding is mating between relatives. The most extreme form of inbreeding is self-fertilization, which occurs naturally in many plants and some invertebrate animals. Because it is relatively simple to treat mathematically, and illustrates most of the important features of less extreme forms of inbreeding, we will consider the population-genetic effects of selfing in some detail.

If we begin by considering a population that consists exclusively of self-fertile heterozygotes of genotype Aa . The initial frequency of the A allele (p) is therefore 0.5. After one generation of selfing, the population would consist of $1/4$ AA , $1/2$ Aa , and $1/4$ aa individuals. So the proportion of heterozygotes (the heterozygosity wrt this one locus) has gone from 1 to $1/2$ in a single generation of self-fertilization. However, the allele frequency remains $p=0.5$. In the next generation, only heterozygotes can produce heterozygous offspring, and only half the offspring of any heterozygote will also be heterozygotes. So the heterozygosity after two generations of selfing is now $1/2 \times 1/2 = 1/4$, and the frequency of the two homozygous genotypes is $3/8$. Three generations of selfing reduce the heterozygosity to $1/8$, and so forth.

So, the primary effect of selfing is that it reduces the heterozygosity by $1/2$ every generation. In particular, note that the allele frequency does not change, only the proportion of heterozygotes. Non-

random mating therefore causes deviations from Hardy-Weinberg proportions, but does not change allele frequencies.

In weaker forms of inbreeding, such as matings between cousins, the same thing happens, only more slowly. A convenient measure of the effect of inbreeding is based on the reduction in heterozygosity that occurs.

$$F = 1 - (\text{observed Het/expected Het}) \text{ so}$$
$$F = (1 - H/2pq) \text{ and}$$
$$H = 2pq (1 - F)$$

7.5.2 INBREEDING IS GENERALLY HARMFUL

We saw above that inbreeding decreases heterozygosity. The obvious corollary is that inbreeding also increases homozygosity, the proportion of individuals within a population that are homozygous with respect to specified loci. In most species, inbreeding is harmful, and much of the effect is due to homozygosity for rare deleterious recessive alleles. Remember that a recessive allele does not affect the phenotype unless it is homozygous, and that rare alleles are usually present in heterozygotes, not in homozygotes, since the frequency of the rare (q is small) homozygote is given by q^2 in a random-mating population. However, with inbreeding, the proportion of homozygotes increases. In a completely inbred population, the proportion of homozygotes for the rare allele would approach q , instead of q^2 . If a rare recessive deleterious allele has a frequency of $q=0.1$ in a random-mating population, the proportion of individuals homozygous for the deleterious allele is $q^2=0.01$. In completely inbred population ($F=1$), the proportion of individuals homozygous for the deleterious allele is $q=0.1$. So there are 10 times as many individuals expressing the deleterious phenotype in a completely inbred population.

In humans, the most common type of close inbreeding is between first cousins ($F=1/16$). Such matings result in an increase in frequency of genotypes that are homozygous for rare, recessive alleles. The increase in frequency of homozygotes for such an allele can be calculated by noting that, in inbreeding populations, the frequency of homozygotes for the less common allele is given by the formula:
 $q^2(1-F)+qF$

[Since, in an inbred population, the proportion of heterozygotes is reduced by $2pqF$ compared to a non-inbred population. Of this "missing" $2pqF$ of the population half are homozygotes for one allele and half are homozygotes for the other. Therefore the proportion of homozygotes for the rarer allele becomes $q^2 + pqF$. Substituting $(1-q)$ for p , we get $q^2 + (1-q)qF = q^2 + qF - q^2F = q^2(1-F) + qF$.]

Example: The frequency of albinism (homozygous for aa) in American Caucasians is 1 in 20,000 ($q^2=0.00005$). But the frequency among offspring of first-cousin marriages is $[0.00005*(15/16)+0.007*(1/16)] = 0.0005$ or 1 in 2000.

Inbreeding also has implications for species with small population sizes, for species in zoos, or for domestic species of plants and animals that are often inbred to "fix" some desirable trait. Highly inbred strains are usually all fixed for the same allele at most loci. This is because, as inbreeding eliminates heterozygotes, genetic drift within a single strain will tend to eliminate alternate alleles. The result is a strain where each individual is highly homozygous, and different individuals are genetically identical. Such populations suffer from inbreeding depression, and suffer the long-term deleterious consequences of the lack of genetic variation within a population--little ability to evolve disease resistance to new pathogens, or to adapt to changing environmental conditions.

7.6 RANDOM GENETIC DRIFT

Since populations are not infinite in size, allele frequencies can change due to chance. If only two individuals in a population reproduce, one may by chance carry a rare allele. So the allele frequency would go from very rare in one generation to 1/4 in the next. A process of sampling takes place every generation, in which the very large number of gametes produced by individuals in a population result in the production of a much smaller number of actual offspring.

Figure 19.11 in your text illustrates the effects of genetic drift in populations that contain only 20 diploid individuals each generation. In generation 0, all 20 individuals have one copy of allele A and one copy of allele a. A computer program was then written to simulate the effects of random mating and Mendelian segregation. The table reports the number of A alleles within each population over a number of generations. After 30 generations, two of the four populations have become fixed for A or for the alternate allele a.

Random genetic drift causes random changes in allele frequencies over time. It does not generally cause deviations from HW equilibrium, as does inbreeding. So the main effect of drift is changes in allele frequencies.

Even large populations, over a long period of time, can experience drastic changes in gene frequencies due to random genetic drift. If random genetic drift is the only force operating within a population, all loci will eventually become fixed for a single allele. This means there is no genetic variation at all left within such a population. This is why small population sizes in species of organisms that are threatened or endangered is a cause of concern. Without a reserve of genetic variation, a population will be unable to respond genetically to changing environmental conditions, to new diseases or parasites, or to the introduction of new competitors, predators or resources. The relatively new field of Conservation Genetics has grown from the need for biologists and wildlife managers to understand the forces affecting genetic variation in small populations.

Several factors can retard the rate of loss of genetic variation due to genetic drift:

1. Large population size
2. migration between subpopulations
3. mutation
4. natural selection ('balancing' selection such as frequency-dependent selection, environment-dependent fitness, and heterozygote superiority)

7.7 GENE FLOW (MIGRATION)

Populations often exchange migrants. This leads to gene flow between different populations of the same species. Sometimes, these different populations will have different alleles or different allele frequencies. These differences can arise due to selection (i.e, the populations may be adapted to different environments), or due to random genetic drift. In general, the effect of gene flow is to equalize allele frequencies among populations that exchange migrants, even if these populations have different allele frequencies because they are adapted to different environments. So migration can work against selection, and lead to non-optimal allele frequencies. Migration, like random genetic drift, can lead to non-adaptive evolution.

For example, DDT resistance may evolve in a single population of insect pest that is exposed to DDT. The alleles conferring this resistance can then spread to other populations (even populations that have not yet been exposed to DDT) by gene flow.

The simplest type of gene flow takes place when an allele enters a population from an outside source. An island model of migration occurs when smaller, isolated populations (as might occur on islands) receive migrants from a large (continental) population. In this scenario, gene flow occurs in only one direction.

In a metapopulation (group of populations connected by gene flow) in which the continental and island population have different allele frequencies, the process of migration can be modelled as follows. Let the proportion of migrants moving into island population each generation be m and the proportion of non-migrants be $(1-m)$. Before migration, let the frequency of the A_2 allele on an island be q , and the allele frequency among migrants coming to the island be q_m . After a single generation of migration, the allele frequency on the island is:

$$q' = (1-m)q + mq_m = q - m(q - q_m)$$

The change in allele frequency after one generation of migration is

$$\Delta q = q' - q = -m(q - q_m)$$

The allele frequency on the island reaches an equilibrium when $\Delta q = 0$. That is, when $q = q_m$. So, the allele frequency on the island gets closer to that on the continent in every generation, until the two frequencies are equal. Intuitively, this makes sense, since the gene flow is occurring in only one direction.

So, in the face of one-way migration, the allele frequencies on an island will approach that on the mainland even if a different allele frequency is favored by selection. If selection and migration are operating in different directions, the island population will eventually reach migration-selection equilibrium, with allele frequencies in between that favored by selection, and that which would result only from the pressure of migration.

In a set of island populations, if genetic drift is operating in addition to selection and migration, the mean allele frequency among the islands will equal that of migration-selection equilibrium. But any individual island will also experience random deviations from this mean, due to genetic drift.