

Lecture 26

In this lecture we will consider how allele frequencies can change under the influence of mutation and selection.

The first consider the conversion of a wild type gene to an altered allele by mutation:

μ
 $A \rightarrow a$ μ =mutation rate (probability of a mutation/generation)

$$\Delta q_{\text{mut}} = \mu f(A) = \mu p \approx \mu$$

Typical mutation rates vary from $\mu = 10^{-4} - 10^{-8}$

Thus, in the absence of any other effects, such as selection, for any given gene the frequency of mutant alleles will increase a little each generation because of new mutations

Consider the disease phenylketonuria (PKU), which is an autosomal recessive defect in the enzyme phenylalanine hydroxylase. The absence of the enzyme prevents phenylalanine from being metabolized causing unusually high levels of phenylalanine in the body leading to severe mental retardation.

Say, that for PKU, $\mu = 10^{-4}$. The frequency of PKU will then slowly increase each generation.

When the allele frequency gets high enough selection against homozygotes will counterbalance new mutations and q will stay constant. In order to treat selection quantitatively we need an additional concept.

S = selective disadvantage; and fitness = $1-S$

If a genotype has $S = 0.75$ then fitness = 0.25, meaning that individuals with this genotype will reproduce at a rate of only 25% relative to an average individual. Fitness can be thought of as a combination of survival and fertility.

Recall that for alleles in H-W equilibrium (random mating) the genotype frequencies will be:

$$f(A/A) = p^2, f(A/a) = 2pq, f(a/a) = q^2$$

Genotype	frequency	after selection	Δ frequency
A/A	p^2	p^2	0
A/a	$2pq$	$2pq$	0
a/a	q^2	$q^2(1 - S)$	$-Sq^2$

$$\Delta q_{sel} = -Sq^2$$

In the steady state: $\Delta q_{sel} + \Delta q_{mut} = 0$, $-Sq^2 + \mu = 0$, $\mu = Sq^2$

$$q = \sqrt{\mu/S}$$

For PKU, q is 10^{-2} and during human evolution $S \approx 1$. Therefore, the estimated value of μ is about 10^{-4} . The actual mutation frequency is probably not this high - and the relatively high q for PKU is probably due to a founder effect in the European population or a balanced polymorphism (see below).

In modern times PKU can be treated by a low-phenylalanine diet so $S < 1$. So the frequency of PKU should start to rise at a rate $\Delta q_{mut} = 10^{-4}$.

Thus, q will only increase by a factor of 1% per generation and it will take a long time for this change in environment to have a significant effect on disease frequency.

Now let's determine the steady state allele frequency for a **dominant** disease with allele frequency $q = f(A)$. In contrast to the situation for recessive alleles, for dominant alleles selection will operate against heterozygotes.

Note that for a rare dominant trait almost all affected individuals are heterozygotes. $q = f(A/A) + 1/2 f(A/a) \approx 1/2 f(A/a)$

Genotype	frequency	after selection	Δ frequency
A/A	-	-	-
A/a	$2pq \approx 2q$	$(1 - S) 2q$	$-2Sq$
a/a	p^2	p^2	0

$$\Delta q_{sel} = 1/2 [\Delta f(A/A)] = 1/2 (-2Sq)$$

$$= -Sq$$

(After selection, $2Sq$ heterozygotes are lost each generation but only 1/2 of their alleles are A. So the net reduction in $f(A)$ is $-Sq$.)

In the steady state: $\Delta q_{sel} + \Delta q_{mut} = 0$, $-Sq + \mu = 0$, $\mu = Sq$

$$q = \mu/S$$

$$\text{For } S = 1, q = \mu$$

In other words, for dominant mutations with fitness = 0, the only instances of the disease will be due to new mutations. This makes sense because mutant alleles cannot be passed from one generation to the next. In this case, the number of affected individuals will be 2μ .

When $S < 1$ the frequency can get quite high. A good example of this is Huntington's disease which has a late onset of degeneration of neuromuscular system at > 35 yrs. This disease is bad personally but doesn't decrease reproductive fitness much.

For the final example of a balance between mutation and selection, consider an **X-linked recessive** allele with frequency $q = f(a)$. For rare alleles the vast majority of affected individuals who are operated on by selection are males, and new mutations will increase the allele frequency $\Delta q_{mut} \approx \mu$

Genotype	frequency	after selection	Δ frequency
$X^A Y$	p	p	0
$X^a Y$	q	$(1 - S)q$	$-Sq$

Note that in a population of equal numbers of males and females, 1/3 of the X chromosomes will be in males.

Therefore,

$$\begin{aligned}\Delta q_{\text{sel}} &= \frac{1}{3} [\Delta f(\text{X}^{\text{a}} \text{Y})] = \frac{1}{3} (-Sq) \\ &= -Sq/3\end{aligned}$$

In the steady state: $\Delta q_{\text{sel}} + \Delta q_{\text{mut}} = 0$, $-Sq/3 + \mu = 0$, $\mu = Sq/3$

$$q = 3\mu/S \qquad \text{For } S = 1, q = 3\mu$$

For X-linked recessive mutations with fitness = 0, exactly one third of the alleles in a population will be new mutations. This relationship has been demonstrated for the debilitating X-linked diseases hemophilia A and Duchenne muscular dystrophy.

Balanced Polymorphism

Now we will consider a situation in which an allele is deleterious in the homozygous state but is beneficial in the heterozygous state. The steady state value of μ will be set by a balance between selection for the heterozygote and selection against the homozygote.

We will need a new parameter that represents the increased reproductive fitness of heterozygote over an average individual.

h = heterozygote advantage

Genotype	frequency	after selection	Δ frequency
A/A	p^2	p^2	0
A/a	$2pq \approx 2q$	$(1 + h) 2q$	$2hq$
a/a	q^2	$(1 - S)q^2$	$-Sq^2$

$$\begin{aligned}\Delta q &= \Delta f(\text{a/a}) + \frac{1}{2} \Delta f(\text{A/a}) = -Sq^2 + \frac{1}{2}(2hq) \\ &= -Sq^2 + hq\end{aligned}$$

Say $S = 1$, then $\Delta q = 0$ when $q^2 = hq$ i.e. $h = q$

The possibility of a subtle selection for (or against) the heterozygote for an allele that appears to be recessive means that in practice the estimates of μ from allele frequencies are quite unreliable.

For example, $q = 10^{-2}$. This could mean $\mu = 10^{-4}$ and $h = 0$ or $\mu < 10^{-4}$ and $h = 10^{-2}$. Since a 1% increase in heterozygote advantage would be essentially unmeasurable we couldn't distinguish these possibilities.

The best understood case of balanced polymorphism is **sickle-cell anemia**

The allele of hemoglobin known as Hb^S is recessive for the disease but is dominant for malarial resistance. Hb^S is most prevalent in a number of different equatorial populations where malaria is common: sub-Saharan Africa, the Mediterranean, and Southeast Asia.

In parts of Africa the frequency of the disease can be as high as $\sim 2.6\%$, which means that in these populations $q = 0.16$.

During human history sickle cell disease would almost certainly be fatal thus $S \approx 1$ and therefore h must have been about 0.16. This indicates that during evolution the reproductive advantage for an Hb^S heterozygote is 16%.

Many of the most prevalent genetic diseases are suspected to be at a relatively high frequency because of balanced polymorphism.

Cystic Fibrosis: Autosomal recessive mutations in CFTR (Cystic fibrosis transmembrane conductance regulator). Mutants disrupt Cl^- transport leading to disturbed osmotic balance across in epithelial cell layers of the lungs and intestine.

Incidence in European populations $\approx 1/2000$. Thus, $q = 0.05$

This high frequency is probably not due to either high mutation frequency or founder effect (many different alleles have been found although 70% are $\Delta F508$).

The hypothesis is that heterozygotes may be more resistant to bacterial infections that cause diarrhea such as typhoid or cholera and that this selection was imposed in densely populated European cities.

A second example is a set of different autosomal recessive **lysosomal storage disorders**

Disease	Enzyme	Allele frequency (maximum)
Gaucher	glucocerebrosidase	0.03
Tay-Sachs	hexosaminidase A	0.017
Nieman-Pick	sphingomyelinase	0.01

All three enzymes are involved in breakdown of glycolipids in the lysosome. When these enzymes are defective (in individuals heterozygous for the disease allele) excessive quantities of glycolipids build up in cells and can have pathological effects. In particular all three diseases are characterized by mental retardation because of excess glycolipids in neurons.

All three diseases are ~ 100x more common in Ashkenazi Jewish populations. This group arrived in central Europe in 9th century AD and is currently distributed among US, Israel, and the former Soviet Union. The competing theories to explain the unusually high allele frequencies are balanced polymorphism or founder effect.

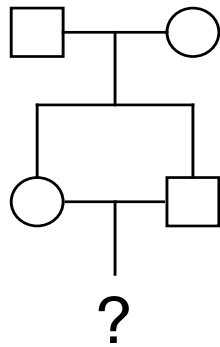
Lecture 27

Effects of Inbreeding:

Today we will examine how inbreeding between close relatives (also known as consanguineous matings) influences the appearance of autosomal recessive traits.

Note that inbreeding will not make a difference for dominant traits because they need only be inherited from one parent or for X-linked traits since they are inherited from the mother.

Consider an extreme case of inbreeding namely a brother-sister mating.



A useful concept is the **Inbreeding Coefficient** = **F** which is defined as the likelihood of homozygosity by descent at a given locus.

If we consider a locus with different alleles in each grandparent: **A1**, **A2**, **A3**, **A4**,

F is the probability that the grandchild will be either **A1/A1**, **A2/A2**, **A3/A3**, **A4/A4**

$$p(\mathbf{A1/A1}) = 1/2 \cdot 1/2 \cdot 1/4 = 1/16$$

$$p(\mathbf{A2/A2}) = \quad \quad \quad " \quad \quad = 1/16$$

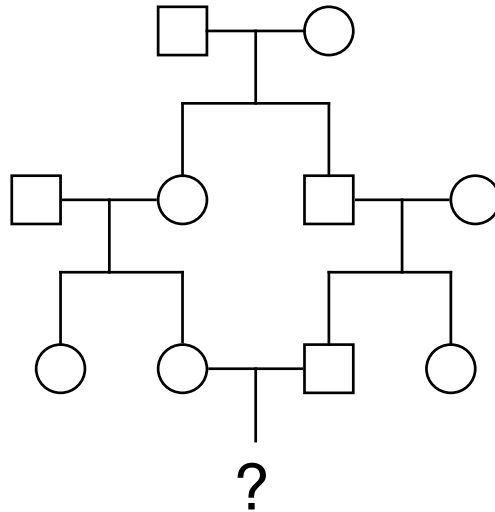
$$p(\mathbf{A3/A3}) = \quad \quad \quad " \quad \quad = 1/16$$

$$p(\mathbf{A4/A4}) = \quad \quad \quad " \quad \quad = 1/16$$

$$p(\text{homozygous by descent}) = 4 \cdot 1/16 \quad \quad \quad \mathbf{F} = 1/4$$

A brother-sister mating is the simplest case but is of little practical consequence in human population genetics since all cultures have strong taboos against this type of consanguineous mating and the frequency is extremely low.

However, 1st cousin marriages do happen at an appreciable frequency. Let's calculate **F** for offspring of 1st cousins.



$$p(A1/A1) = 1/2 \cdot 1/2 \cdot 1/2 \cdot 1/2 \cdot 1/4 = 1/64$$

$$p(A2/A2) = \quad \quad \quad " \quad \quad \quad = 1/64$$

$$p(A3/A3) = \quad \quad \quad " \quad \quad \quad = 1/64$$

$$p(A4/A4) = \quad \quad \quad " \quad \quad \quad = 1/64$$

$$p(\text{homozygous by descent}) = 4 \cdot 1/64, \quad \mathbf{F} \text{ for 1}^{\text{st}} \text{ cousins} = 1/16$$

Consider a rare recessive allele **a** at frequency $f(a) = q = 10^{-4}$

For random mating the frequency of homozygotes is $f(a/a) = q^2 = 10^{-8}$

Imagine a hypothetical situation where only 1st cousins mated. In that case the frequency of homozygotes would be:

$$f(a/a) = p(\text{homozygous by descent}) \times p(\text{allele is } a)$$

$$= \quad \quad \quad \mathbf{F} \quad \quad \quad \times \quad \quad \quad \mathbf{q}$$

$$f(a/a) = 1/16 \times q = 6.3 \times 10^{-6}$$

Thus there would be 600 times more affected individuals for 1st cousin matings than for random mating. But 1st cousin marriages are rare and their actual impact on the frequency of homozygotes in a population will depend on the frequency of 1st cousin marriages.

In the U.S. the frequency of 1st cousin marriages is ≈ 0.001

p (affected because of 1st cousin mating) = $\frac{1}{16} q 10^{-3} = 6.3 \times 10^{-9}$

p (affected because of random mating) = 10^{-8}

Thus, $\sim 1/3$ of affected individuals will come from 1st cousin marriages

Note that this proportion depends on allele frequency such that traits caused by very rare alleles will more often be the result of consanguinity

For rare diseases, it is often quite difficult to tell whether or not they are of genetic origin. A useful method to identify disorders that are likely to be inherited is to ask whether an unusually high proportion of affected individuals have parents that are related to one another.

Now let's consider the problem of recessive lethal mutations in the genome:

We have already seen that the frequencies of recessive, loss of function alleles are usually in the range of 10^{-3} - 10^{-4}

This may seem like a comfortably small number but given that the total number of human genes is about 2×10^4 , each of us must be carrying many recessive alleles. Assuming that about 50% of genes are essential, each person should carry an average of approximately 1-10 recessive lethal mutations!

Genetic Load: lethal equivalents per genome.

Usually the genetic load is not a problem since it is very unlikely that both parents will happen to have lethal mutations in the same genes. However, that chance is considerably increased for parents that are 1st cousins.

As we have already calculated, the probability that a grandparental allele will become homozygous is $1/64$ for 1st cousins

Thus, each recessive lethal allele for which one of the grandparents in a carrier will contribute an increased probability of 0.016 that the grandchild will be homozygous and therefore be afflicted by a lethal inherited defect.

To look for this effect we will use the frequency of stillbirth or neonatal death from 1st cousin marriages. We must also be careful to subtract the background frequency of stillbirths and neonatal deaths that are not due to genetic factors. These frequencies can be obtained from the cases where parents are not related.

	unrelated parents	1st cousins	difference
Observed frequency of still-birth or neonatal death	0.04	0.11	0.07

Average number of recessive lethals in both grandparents = $0.07/0.016 = 4.4$

Thus each grandparent has an average of 2.2 recessive lethal alleles.