Lecture 3A

Population Genetics
Key issues

- More than 2 alleles exist in a population for most genes and these have derived from mutations.

- Each individual has only two alleles for autosomal genes.
Key issues

- Allele frequencies can be used to define a population and predict results of group matings.

- Differences in allele frequencies are a major source of variation between populations.

- Hardy-Weinberg equilibrium can exist in a population for one trait but not for another.
Predicting Offspring

- Remember our dog example from lecture 2?
- We mated two heterozygous dogs and we expected -
## What pups?

<table>
<thead>
<tr>
<th>Female gametes</th>
<th>B</th>
<th>BB</th>
<th>Bb</th>
</tr>
</thead>
<tbody>
<tr>
<td>black</td>
<td>black</td>
<td>black</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>Bb</th>
<th>bb</th>
</tr>
</thead>
<tbody>
<tr>
<td>black</td>
<td>brown</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What if we had a group of females and males and we didn’t know which ones mated?

What if there are more than 2 alleles for this gene?
Extending our matings

- Many genes have more than 2 alleles in the population.
  - Eg in humans the ABO blood system

- New alleles arise by mutation – could have been thousands of years ago or yesterday!

- A mutation is a change in the DNA – single base, many bases, deletions, insertions or inversions
Many to 2

- No matter how many alleles exist in a population for a particular gene, an individual can only carry 2 alleles

- if they are autosomal or X-linked in a female

- or 1 copy if X-linked in a male (considering mammals here)
Predicting offspring

- What are the implications for predicting offspring?

- What phenotypes and in what proportions can we expect if we do not know the actual parents?

- We need to know the proportions of each allele in the population and use an extended Punnett square.
Hardy-Weinberg Equilibrium

- The proportion of an allele in a particular population is called the allele frequency.

- A population can then be described by the types and proportions of alleles present for a series of genes.
Calculating allele frequencies

Example:

- We have a population of frogs in a lake
- We are interested in the gene, $L$, that directs the ‘loudness’ of their call – loud or soft.
- Collect all the sperm – 1,000,000 – in a bucket
- Examine with a magic microscope and divide into $L$ or $l$ sperm.
Allele frequencies -2

- There are 700,000 sperm of the $L$ type and 300,000 of the $l$ type.

- The frequencies are
  
  \[
  f(L) = \frac{700,000}{1,000,000} = 0.7 = p \\
  F(l) = \frac{300,000}{1,000,000} = 0.3 = q
  \]
Allele frequencies

- We can do the same with the eggs we collect from the females

- \( f(L) = \frac{35,000}{50,000} = 0.7 = p \)

- \( F(l) = \frac{15,000}{50,000} = 0.3 = q \)
Generalisations

1. We can see that \(0.7 + 0.3 = 1\)
   Or \(p + q = 1\)
   If there are 3 alleles then \(p + q + r = 1\)

2. We had a very large sample – 1,000 males and 1,000 females and the eggs and sperm mixed randomly. That is we can say we had a large random mating population with respect to this trait.
3. The allele frequencies are the same in males and females.

4. We can use these frequencies to predict the possible offspring with respect to this trait using a Punnett square and multiplying the allele frequencies.

<table>
<thead>
<tr>
<th></th>
<th>0.7 L</th>
<th>0.3 l</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7 L</td>
<td>0.49 LL</td>
<td>0.21 Ll</td>
</tr>
<tr>
<td>0.3 l</td>
<td>0.21 Ll</td>
<td>0.09 ll</td>
</tr>
</tbody>
</table>
Generalisations - 3

- We expect

\[ 0.49 \, LL \quad 0.42 \, Ll \quad 0.09 \, ll \]

And if Loud call \( L \) is dominant to soft call, \( l \)

\[ 0.91 \, \text{Loud calling frogs} : 0.09 \, \text{soft calling frogs} \]

If we substitute \( p \)’s and \( q \)’s for the allele frequencies

\[ p^2 \, LL \quad 2pq \, Ll \quad q^2 \, ll \]
5. There were no mutations occurring

6. and no migration of individuals into or out of the lake.

7. There was no selection by either sex or the environment for or against either of the genotypes or alleles.
Hardy-Weinberg Equilibrium

- If these generalisations are met then the population is said to be in Hardy-Weinberg equilibrium.

In a large, randomly mating population free from evolutionary forces, the allelic frequencies do not change and the genotypic frequencies stabilise after 1 generation of mating. In the case of 2 alleles, A and a, with frequencies $p$ and $q$, the genotypic frequencies at equilibrium are $p^2$, $2pq$ and $q^2$ for AA, Aa and aa respectively.
Real-life calculations

- Assuming 1 gene with 2 alleles
  \[ A_1 \text{ and } A_2 \]

  The dominance relationship between these alleles will vary
1. The Add-up method

- We use this if we know the numbers of individuals for ALL genotypes.

- This is the most accurate of methods

- It does not depend on the population being in Hardy-Weinberg equilibrium
Example: Herd of 100 Shorthorn cattle

<table>
<thead>
<tr>
<th>genotypes</th>
<th>Red</th>
<th>Roan</th>
<th>White</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR</td>
<td>30</td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Rr</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rr</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example: Herd of 100 Shorthorn cattle

<table>
<thead>
<tr>
<th>genotypes</th>
<th>Red</th>
<th>Roan</th>
<th>White</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR</td>
<td>30</td>
<td>50</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>Rr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype frequencies</td>
<td>0.3</td>
<td>0.5</td>
<td>0.2</td>
<td>1</td>
</tr>
</tbody>
</table>
To calculate the allele frequencies we need to add up the actual number of a particular allele.

<table>
<thead>
<tr>
<th></th>
<th>RR</th>
<th>Rr</th>
<th>rr</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>50</td>
<td>20</td>
</tr>
</tbody>
</table>

\[ f(R) = p = \frac{2(30) + 50}{200} = 0.55 \]

\[ f(r) = q = \frac{2(20) + 50}{200} = 0.45 \]

or \[ 1 - 0.55 = 0.45 = q \]
2. The square-root method

- We use this if we can’t tell the heterozygotes from one of the homozygotes.

- We have to assume the population is in H-W equilibrium.

- That is we assume the 
  \[ f(\text{homozygous recessives}) = q^2 \]
Example:

In a population of dogs we have 16% brown dogs compared to black ones. What are the observed phenotypic, allele and expected genotypic frequencies?

a) The observed phenotypic frequency of brown = 16% = 0.16 so

Frequency of black = 84% = 0.84
To calculate the allele frequencies:

\[
f(bb) = 0.16 = q^2
\]

So \( f(b) = \sqrt{0.16} = 0.4 = q \)

And \( f(B) = 0.6 = p \)

c) Expected genotypic frequencies

\[
f(BB) = p^2 = (0.6)^2 = 0.36
\]

\[
f(Bb) = 2pq = 2(0.6)(0.4) = 0.48
\]

\[
f(bb) = q^2 = 0.16
\]
**Summarising**

<table>
<thead>
<tr>
<th>Observed Phenotypic frequencies</th>
<th>f(Black) = 0.84; f(brown) = 0.16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele frequencies</td>
<td>f(B) = 0.6</td>
</tr>
<tr>
<td></td>
<td>f(b) = 0.4</td>
</tr>
<tr>
<td>Expected genotypic frequencies</td>
<td>f(BB) = 0.36; f(Bb) = 0.48; f(bb) = 0.16</td>
</tr>
</tbody>
</table>
3. Frequencies for X-linked alleles

- In mammals, males have only 1 copy of an X-linked allele so the proportion of males with the trait directly reflects the allele frequency.

- Example: If 6% of males present with a recessive X-linked disorder then
  \[ f(X^dY) = 0.06 = f(X^d) = q \]
  \[ f(X^DY) = f(X^D) = 0.94 = p \]

- In females with 2 X chromosomes
  \[ f(X^D X^D) = p^2 \]
  \[ f(X^D X^d) = 2pq \]
  \[ f(X^d X^d) = q^2 \]
  \[ = 0.8836 \]
  \[ = 0.1128 \]
  \[ = 0.0036 \]
X-linked method

- X-linked recessive traits are expressed more often in male mammals than in females.

  Eg: in Humans, colourblind women and those with haemophilia are quite rare compared to affected men.
Estimating offspring

- If we know the allele frequencies in males and females in a population we can calculate the expected offspring.

- Works on the multiplication principle by expanding

\[(p + q)^2 = p^2 + 2pq + q^2\]

- If there are more than 2 alleles we just add in more symbols to the basic equation.
Example: 3 coat colours in horses – Chestnut (DD), Palomino (Dd) and Cremello (dd) as seen in a wild herd.

Given the frequencies are \( f(D) = 0.45 \) and \( f(d) = 0.55 \).

We are not told if the frequencies are the same in males and females so assume they are.
To predict the types and proportions in the next generation:

<table>
<thead>
<tr>
<th></th>
<th>0.45 D</th>
<th>0.55 d</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.45 D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.55 d</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Estimating offspring - 3

- To predict the types and proportions in the next generation:

<table>
<thead>
<tr>
<th></th>
<th>0.45 D</th>
<th>0.55 d</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.45 D</td>
<td>0.202 DD</td>
<td>0.248 Dd</td>
</tr>
<tr>
<td>0.55 d</td>
<td>0.248 Dd</td>
<td>0.303 dd</td>
</tr>
</tbody>
</table>

20% Chestnut  50% Palomino  30% Cremello
Estimating offspring - 4

- The frequencies do not have to be the same in males and females.

- Modify the allele frequencies in the sperm and eggs to reflect this.
Example:

A farmer wanted to increase the frequency of the polled allele in his herd of merinos. He imported polled rams and mated them to his ewe flock. Occasionally a polled lamb had been born but only 2% of the time indicating that in the ewes the frequency for the polled allele, $p$, is 0.14.

What would the allele frequencies be in the next generation?

How many polled lambs should be born out of a flock of 500 ewes (if you assume 100% lambing and no mortality)?
As all rams are polled, \( f(p_m) = 1 \)

In females, \( f(p_f) = 0.14 \)

And \( f(P_f) = 0.86 \)

- \( f(P) = 0.86 / 2 = 0.43 \)
- \( f(p) = [0.86 + 2(0.14)] / 2 = 0.57 \)

Note that the allele frequencies will be the same in male and female lambs.

The number of polled lambs = \( 0.14 \times 500 = 70 \)
Small populations

- Using a small number of males can influence the allele frequencies dramatically.

- This has large implications for small populations and threatened species.
Reading pedigree charts
How is this trait inherited?

Keep it simple –
- Consider each type in order
- Put in symbols
- Does it fit the pedigree?
  - AD
  - AR
  - XD
  - XR
  - Cytoplasmic