Mendelian Population Genetics

• Evolution: change in allele frequencies within a population over time
A mouse is a vehicle for mouse gene replication

Mice with Aa genotypes
IF no mutation...Gametes

A gene pool with allele frequencies of 0.6 and 0.4...
**IF random mating**

<table>
<thead>
<tr>
<th>Eggs</th>
<th>Sperm</th>
<th>Zygotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6 A</td>
<td>0.36 AA</td>
<td>0.24 Aa</td>
</tr>
<tr>
<td>0.4 a</td>
<td>0.24 aA</td>
<td>0.16 aa</td>
</tr>
</tbody>
</table>
**Zygotes**

Egg | Sperm | Zygote | Probability
--- | --- | --- | ---
\(A\) & \(A\) | \(AA\) | \(0.6 \times 0.6 = 0.36\)
\(A\) & \(a\) | \(Aa\) | \(0.6 \times 0.4 = 0.24\)
\(a\) & \(A\) | \(aA\) | \(0.4 \times 0.6 = 0.24\)
\(a\) & \(a\) | \(aa\) | \(0.4 \times 0.4 = 0.16\)

Sum = 1

Then zygotes grow up to be adults
IF no survival selection, and IF no biased migration, and IF no biased random death

A population with genotype frequencies of 0.36, 0.48, and 0.16...

THEN you get these adult genotypes, which produce gametes......
Adults produce gametes \textit{IF} no mutation.…

...yields gametes...

\[ A A A A A A \]
\[ A A A A A A \]
\[ A A A A A A \]
\[ A a A a A a \]
\[ A a A a A a \]
\[ A a A a A a \]
\[ a a a a a a \]

...with frequencies of 0.6 and 0.4

\[ A = 0.36 + \frac{1}{2} (0.48) = 0.6 \]
\[ a = \frac{1}{2} (0.48) + 0.16 = 0.4 \]

Those are the same allele frequencies we started with!! \( A = 0.6 \) \( a = 0.4 \)
There was no evolutionary change… IF

• No mutation
• Random mating
• No biased migration
• No biased selected mortality (selection)
• No biased random mortality (drift)
The general case

<table>
<thead>
<tr>
<th>Egg</th>
<th>Sperm</th>
<th>Zygote</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$ &amp; $A_1$</td>
<td>$A_1 A_1$</td>
<td>$p \times p = p^2$</td>
<td></td>
</tr>
<tr>
<td>$A_1$ &amp; $A_2$</td>
<td>$A_1 A_2$</td>
<td>$p \times q = pq$</td>
<td></td>
</tr>
<tr>
<td>$A_2$ &amp; $A_1$</td>
<td>$A_2 A_1$</td>
<td>$q \times p = qp$</td>
<td></td>
</tr>
<tr>
<td>$A_2$ &amp; $A_2$</td>
<td>$A_2 A_2$</td>
<td>$q \times q = q^2$</td>
<td></td>
</tr>
</tbody>
</table>

Developed by Hardy 1908
Mind your p’s and q’s

• If \( p \) = frequency of A1 allele
• And \( q \) = frequency of A2 allele
• And if there are only two alleles in the population at this locus
• Then \( p + q = 1 \)
  – Alternatively \( q = 1 - p \)
Randomly combining gametes in the general case

- Sperm are A1 with probability $p$
- Eggs are A1 with probability $p$
- So A1A1 zygotes are produced with probability of $p \times p = p^2$
And so on….

- A1A1 genotype has frequency of $p^2$
- A1A2 genotype has frequency of $pq$
- A2A1 genotype has frequency of $pq$
- A2A2 genotype has frequency of $q^2$

- But A1A2 is the same as A2A1 = 2 $pq$
- Frequencies must add to 1
  
  $$p^2 + 2pq + q^2 = 1$$
Hardy-Weinberg Equilibrium

- If a simple set of assumptions holds, then the allele frequencies in a population will not change.

- If we symbolize allele frequencies as $p$ and $q$, then genotype frequencies are $p^2$, $2pq$, and $q^2$. 

What are those simple assumptions?

• No *selection*: genotypes survive at equal rates (i.e. no survival selection) and contribute gametes to the next generation equally (i.e. no sexual selection)

• No *mutation*: the alleles we are accounting for stayed the same, none disappeared or were created anew by mutation
What are those simple assumptions?

- No *biased migration*: genotypes do not enter or exit the population non-randomly.
- No biased random events: genotypes do not get zapped by lightning (or whatever) non-randomly; if so, called *genetic drift*.
- *Mating is random by genotype*. Violating this assumption affects genotype frequency, not allele frequency. This is NOT the same thing as sexual selection (unequal gametic contribution of genotypes).
What a bunch of malarcky!

- Obviously those five conditions are almost never met!

- So when those conditions are not met, allele frequencies WILL change, i.e. evolution will occur
Still, what good is that?

- Hardy-Weinberg equilibrium model is a null model
- It is a random expectation given a specific set of assumptions
- Specification of those assumptions allows us to see when and how evolutionary change does occur
And what about the only two alleles nonsense?

For example, if there are three alleles with frequencies \( p_1, p_2, \) and \( p_3 \), such that

\[
p_1 + p_2 + p_3 = 1
\]

then the genotype frequencies are given by

\[
(p_1 + p_2 + p_3)^2 = p_1^2 + p_2^2 + p_3^2
\]

\[
+ 2p_1p_2 + 2p_1p_3 + 2p_2p_3
\]

and the allele frequencies do not change from generation to generation.
Non-equilibrium: violations of Hardy-Weinberg

- Mutation
- Selection
- Migration
- Drift
- Non-random mating
Pop Quiz

• A has frequency of 0.9, what is the frequency of a (if only two alleles)?
• \( p^2 + 2pq + q^2 \ldots = \)
• What are the HW genotypic frequencies?
Hardy-Weinberg recap

• If a set of five assumptions is met:
  – Then, allele frequencies won’t change
  – And,
  – Genotypic frequencies follow from
    • $p^2 + 2pq + q^2 = 1$
Mutation as an evolutionary Force

- Mutation critical for genetic variation
- But can mutation, by itself, have much evolutionary impact on allele frequencies?
A = functional allele; a = mutant knockout

- Mutation from A to a
  - Mutation rate of 1 per 10,000
  - plausible but high
- Mutation from a to A negligible
- Suppose, initially A allele has a frequency of 0.9
What is the effect of mutation?

Initial genotype frequencies

<table>
<thead>
<tr>
<th></th>
<th>AA</th>
<th>Aa</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.81</td>
<td>0.18</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>

Initial allele frequencies

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>0.1</td>
<td></td>
</tr>
</tbody>
</table>

Final genotype frequencies

<table>
<thead>
<tr>
<th></th>
<th>AA</th>
<th>Aa</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.80984</td>
<td>0.18014</td>
<td>0.01002</td>
<td></td>
</tr>
</tbody>
</table>

Final allele frequencies

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.89991</td>
<td>0.10009</td>
<td></td>
</tr>
</tbody>
</table>

Mutation converts copies of A into new copies of a at the rate of 1 per 10,000
Effect over time?

Figure 5.23  Over very long periods of time, mutation can eventually produce appreciable changes in allele frequency
Mutation - selection balance

- Mutations are random
- In finely tuned organisms, most mutations are bound to be bad
- Bad mutations selected against
- Also created anew
- Rate of creation by mutation = rate of removal by selection
- Called mutation-selection balance
Deleterious mutations selected against

- How common should any particular deleterious mutation be?
- Depends on how bad (strength of selection)
- Depends on how often re-created (mutation rate)

\[ \hat{q} = \sqrt{\frac{\mu}{s}} \]
Is that useful information?

• Two examples
  
• Spinal muscular atrophy, study by Wirth et al.
  – Caused by mutation in \textit{telSMN}
  – Telomeric survival motor neuron

• Cystic fibrosis, study by Pier et al.
  – Most common lethal autosomal recessive disease in Caucasians
  – Proximate cause bacteria \textit{Pseudomonas aeruginosa} lung damage
Spinal muscular atrophy

- *telSMN* knock out mutants frequency of 0.01
- Selection coefficient $\sim 0.9$
- Can use equation to estimate what mutation rate would be needed to counter $s = 0.9$, and give $q^*$ of 0.01
- And the calculated answer is $0.9 \times 10^{-4}$ mutations per allele per generation
- Wirth et al. directly measured rate of $1.1 \times 10^{-4}$
- *Mutation-selection balance*
Cystic fibrosis

- Chromosome 7 mutations for protein called Cystic fibrosis transmembrane conductance regulator (CFTR)
- Expressed in lining of lungs and intestines
- In lungs, enables cells to destroy *Pseudomonas aeruginosa*
- Until recently most cistic fibrosis sufferers died before reproducing (i.e. selection very very strong, approaching $s = 1$)
Cystic fibrosis mutation-selection balance?

• Could mutation selection balance maintain mutants at frequency of 0.02 with $s = 1$?
• Yes, but mutation rate would have to be about $4 \times 10^{-4}$
• Observed mutation rate much lower, ca. $6.7 \times 10^{-7}$
Does something else maintain cystic fibrosis causing mutations at that level?

• This is an important question to ask in this case; note that it was not necessary to ask this in the case of Spinal Muscular Atrophy

• The answer could have important medical implications
A possible answer

• What if mutation(s) favored in another context?
• Recall CFTR expressed in intestine as well as lung
• Pier et al hypothesized a function for most common disease causing allele
  – Prevents intestinal infection by *Salmonella typhi*
    • Agent causes Typhoid fever
DeltaF508 allele, cystic fibrosis, and Typhoid fever

- deltaF508 most common cystic fibrosis causing allele
- In humans, homozygotes would get cystic fibrosis
- Could heterozygotes be protected from Typhoid fever?
Mouse study

- Heterozygotes more resistant to *Salmonella typhi*
- Homozygotes almost completely protected
Pleiotropy and genetic correlation

- deltaF508 single allele with multiple (pleiotropic) effects

- Genetic correlation
  - Lung lining phenotype will show some similarity to intestine lining phenotype
  - Not completely separate traits for selection to act upon independently

- And that’s what is next: selection
Selection as an evolutionary force

- Hardy-Weinberg null model assumptions
- Individuals of all genotypes survive at equal rates, and contribute gametes equally to the gene pool

- Bottom line is differential reproductive success: some genotypes out reproduce others
example
Result

• 7.5% increase in frequency of B1 allele
  7.5% decrease in frequency of B2 allele
• Population has evolved
• Also true for much weaker selection
Selection across generations

Selection scheme

<table>
<thead>
<tr>
<th></th>
<th>$B_1B_1$</th>
<th>$B_1B_2$</th>
<th>$B_2B_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>100</td>
<td>90.0</td>
<td>80.0</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>98.0</td>
<td>96.0</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>99.0</td>
<td>98.0</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>99.5</td>
<td>99.0</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>99.8</td>
<td>99.6</td>
</tr>
<tr>
<td>Weak</td>
<td>100</td>
<td>99.0</td>
<td>99.0</td>
</tr>
</tbody>
</table>

Frequency of allele $B_1$ vs. generation.
Empirical Example: *Adh*

- Alcohol dehydrogenase
  - Enzyme breaks down ethanol
- Cavener & Clegg *Drosophila melanogaster* experiment
  - Two replicate populations on ethanol spiked food
  - Two replicate populations on ethanol free food
- Tracked allele frequency across time
Adh semi-natural selection experiment
Selection and genotypic frequency

• If a set of five assumptions is met:
  – Then, allele frequencies won’t change
  – And,
  – Genotypic frequencies follow from
    • \( p^2 + 2pq + q^2 = 1 \)
• Does selection alter genotype frequency?
  – Sometimes, for one generation
    • Heterozygote advantage
Selection and genotype frequency

Allele freq. 0.5, 0.5
Heterozygotes fitter
After selection
Allele freq. 0.5, 0.5
Predicts genotype
Frequency of
0.25, 0.5, 0.25
Observe 0.167, 0.667, 0.167
Generalizing Selection

• Recall development of Hardy-Weinberg argument
• We took special case of $A = 0.6$ and $a = 0.4$
• Then generalized using $p$ as frequency of the $A$ allele and $q$ as frequency of the alternative allele $a$
• Can we develop a general model of selection?
Change in allele frequency depends on *genotypic* fitness

- Why?, because selection acts on phenotypes not alleles
- Fitness symbolized as $w$
- Fitness of a genotype symbolized with subscripts, so that
- Fitness of $A_1A_1$ genotype has fitness $w_{11}$
- Fitness of $A_1A_2$ genotype has fitness $w_{12}$
- Fitness of $A_2A_2$ genotype has fitness $w_{22}$
How will allele frequency change with unequal genotypic fitness?

• Unequal fitness needs a standard
• Fitness of a genotype is expressed relative to a standard of the \textit{average} fitness of all genotypes

$$\bar{w} = p^2 w_{11} + 2pq w_{12} + q^2 w_{22}$$
Genotypic frequency change

- Genotypic frequency change depends on the fitness of the genotype *relative* to the average fitness.
How have allele frequencies changed?

- How many A1 gametes produced?
- A1 gametes in proportion to A1A1 genotype frequency multiplied by relative genotype fitness
- Plus 1/2 of A1A2 genotype frequency multiplied by relative genotype fitness

\[
p^2w_{11} + pqw_{12} \over \bar{w}
\]

1/2 of 2pq = pq
Same for A2 alleles

- A2 gametes proportional to frequency of A2A2 genotype (q2) multiplied by the relative fitness of the q2 genotype
- Plus 1/2 of the alleles contributed by heterozygotes
  
  that is, 1/2 of the frequency of heterozygotes (2pq) multiplied by heterozygote relative fitness

\[
\frac{pqw_{12} + q^2w_{22}}{w}
\]

1/2 of 2pq = pq
Allele frequency change

- Change = new allele frequency - old allele frequency

\[
\Delta p = \frac{p^2 w_{11} + pqw_{12}}{w} - p
\]

\[
= \frac{p^2 w_{11} + pqw_{12}}{w} - \frac{pw}{w}
\]

\[
= \frac{p^2 w_{11} + pqw_{12} - pw}{w}
\]

\[
= p \left( \frac{pw_{11} + qw_{12} - \bar{w}}{w} \right)
\]
Ta Da!!!

• We have made a general expression that quantitatively describes how allele frequency will change depending on the average fitness of that allele.

• Similarly,

\[
\Delta q = \frac{pq w_{12} + q^2 w_{22}}{\bar{w}} - q
\]

\[
= \frac{q}{\bar{w}} (pw_{12} + qw_{22} - \bar{w})
\]
Patterns of Selection

• Selection on dominant versus recessive alleles
• Selection on heterozygotes and homozygotes
• Frequency dependent selection
Dominant and recessive alleles

- A dominant to a if
- AA genotype = A phenotype
- AND
- Aa genotype = A phenotype
- but
- aa genotype = a phenotype
- If so, A called dominant, a called recessive
Suppose homozygous recessive is lethal

- $w_{aa} = 0$
- How will frequency of $a$ change over time?
  - If $a$ is common?
  - If $a$ is rare?
- If $a$ is common then genotype $aa$ is pretty common
  - If frequency of $a = q$ then genotype frequency $= q^2$
What do we expect

• If at first a is common, it will decrease rapidly
• As it becomes more rare, it will decrease more slowly
  – Because when rare, there are few aa genotypes with fitness = 0
Data from flour beetles

- Dawson started experiment with heterozygotes, so $p = q = 0.5$
- $w_{aa} = 0$
Data agree with qualitative prediction

- Rapid change when aa genotype common
- Slow change when aa genotype rare
Selection against recessive

(a) Selection against a recessive allele and for a dominant allele

Fraction surviving:

\[
\begin{array}{ccc}
AA & Aa & aa \\
1.0 & 1.0 & 0.5 \\
\end{array}
\]

![Graph showing the frequency of a recessive allele and fitness over generations.](image-url)
Selection for recessive
Heterozygotes

- Heterozygotes don’t always express dominant phenotype
- Heterozygotes can have fitness
  - Intermediate
    - Changes rate of evolution, not outcome
  - Higher
    - Heterozygote superiority or overdominance
  - Lower
    - Heterozygote inferiority or underdominance
Overdominance

• Suppose genotypic fitness is as follows:
  \[ w_{AA} = 0.735 \]
  \[ w_{Aa} = 1 \]
  \[ w_{aa} = 0 \]

• Equilibrium frequency of a allele > 0

• Example seen in *Drosophila melanogaster* study of Muaki and Burdick (1958)

• Specific gene effects unknown
Four populations, 2 replicates
Overdominance generalized

- Mean fitness across genotypes

(b) Mean fitness as a function of \( p \) for overdominance
underdominance

- What if heterozygotes have lower fitness?
- Considerably more complicated
- Outcome depends on starting conditions
- One allele goes to fixation (100%)
  - Other allele lost
Underdominance generalized

(c) Mean fitness as a function of $p$ for underdominance

Adaptive landscape
Heterozygote selection outcomes

s and t are selection terms for homozygotes

Both negative = heterozygote superiority (overdominance)

Both positive = heterozygote inferiority (underdominance)
Frequency dependent selection

• Strength of selection depends on frequency
• For example, if the allele that is rare is favored
• Two examples
  – Book scale eating fish
  – Sex
Scale eating fish *Perissodus microlepis*
Lake Tanganyika
- Mouth twisted to side, bites scales off other fish
- Left mouth bites right side of victim; Right mouth bites left; Victim fish wary
Sex

• Sex ratio (males:females) usually about 1:1
• Imagine the fitness of a male if males were really rare
  – That rare male would make lots of babies
• Female fitness if females rare?
  – Most males would not reproduce
• Imagine male female controlled by single locus with two alleles