Lecture 3

Population Genetics, Epidemiology and Evolution

Recap from last time…

• Evolution?
• Selection?
Evolution

• Descent with modification
• Change in gene frequency in a population over time
  – Mutation
  – Migration
  – Selection
  – Genetic Drift
  – Genetic Drift

Measures of Evolution

• How do we know evolution is happening?
• Population genetics
  – Evolution + Genetics
  – Darwin + Mendel
  – How populations change genetically over time
Modern Synthesis

• Comprehensive theory of evolution
  – Fisher - inheritance of Mendelian traits
  – Haldane - natural selection
  – Dobzhansky and Wright - genetics
  – Mayr - biogeography
  – Simpson – paleontology (rate of $\Delta$)
  – Stebbins – variation in plants

• Huxley
  – “Evolution: The Modern Synthesis”

Modern Synthesis

• Genetic variation in populations arises by chance through mutation and recombination
• Evolution consists primarily of changes in the frequencies of alleles between one generation and another as a result of
  – genetic drift
  – gene flow
  – natural selection
• Speciation occurs gradually when populations are reproductively isolated
Mendelian “Review”

- Genotype/Phenotype
- Dominant/Recessive
- Heterozygous/Homozygous
- Punnet Squares

Each true-breeding plant of the parental generation has identical alleles, PP or pp. Gametes (circles) each contain only one allele for the flower-color gene. In this case, every gamete produced by one parent has the same allele.

Union of the parental gametes produces F₁ hybrid having a Pp combination. Because the purple-flower allele is dominant, all these hybrids have purple flowers.

When the hybrid plants produce gametes, the two alleles segregate; half the gametes receive the P allele and the other half the p allele.

This box, a Punnett square, shows all possible combinations of alleles in offspring that result from an F₁, Pp x Pp, cross. Each square represents an equally probable product of fertilization. For example, the bottom left box shows the genotypic combination resulting from a p egg fertilized by a P sperm.

Random combination of the gametes results in the 3:1 ratio that Mendel observed in the F₂ generation.
Mendelian “Review”

• Aa x AA ->
  – A. 1 AA : 2 Aa : 1 aa
  – B. 1 AA : 3 Aa
  – C. 1 AA
  – D. 1 aa : 2 Aa
  – E. 1 Aa : 1 AA

Predicting Genotypes

• For a particular locus with alleles A and a
• Frequency of A = p and a = q
• We know that p + q = 1
• If there is random mating
  – Frequency of AA = p * p = p²
  – Frequency of aa = q * q = q²
  – Frequency of Aa = pq + pq = 2pq
• We also know then that
  – p² + q² + 2pq = 1
Hardy Weinberg

• Populations with consistent proportions for p and q as follows
  – Homozygous Dominant = $p^2$
  – Homozygous Recessive = $q^2$
  – Heterozygous = $2pq$

• Are said to be at Hardy Weinberg Equilibrium

Hardy Weinberg

• Allele and genotype frequencies

[Diagram showing genetic inheritance through generations, including allele and genotype frequencies.]
Hardy Weinberg Problem I

- Sickle Cell Anemia
  - SS = susceptible to malaria but no SCA
  - ss = non-susceptible but SCA \(\rightarrow\) mortality
  - Ss = non-susceptible and no SCA

- What do we expect proportions of ss? Ss?
- E.g. 4% ss - what are proportions of Ss
  - \(0.04 = ss = q^2 \rightarrow q = \sqrt{0.04} = 0.2\)
  - \(p = 1 - q = 1 - 0.2 = 0.8\)
  - \(Ss = 2pq = 2(0.8)(0.2) = 0.32\)

Hardy Weinberg Problem II

- What would it take to increase Ss proportion to 50% (from 32%)?
  - \(Ss = 2pq = 0.5 \rightarrow pq = 0.25\)
  - \(q = 1 - p \rightarrow p (1-p) = 0.25 \rightarrow p = 0.5\)
  - Which means
    - \(aa\) goes to 0.25 from 0.04, over 6x
Hardy Weinberg Problem III

• How many of you can roll your tongues?
  – A - Yes
  – B - No

• What is the percentage of heterozygous tongue-rollers?
  – Yes = $p^2 + 2pq$
  – No = $q^2$
  – $q = \sqrt{No}$
  – $p = 1 - q$
  – Heterozygous = $2pq$

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tongue Rollers</td>
<td>7</td>
<td>0.77777777</td>
</tr>
<tr>
<td>Non-Rollers</td>
<td>2</td>
<td>0.22222222</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

\[ p = 0.5083995479 \]
\[ q = 0.4916004521 \]

\[ \text{Heterozygotes} = 0.4983645377 \]
Conditions for HW Equilibrium

- Extremely Large Population Size
  - No genetic drift
- No Gene Flow
  - No immigration or emigration
- No Mutations
- Random Mating
  - No selecting particular mates
- No Natural Selection

Detecting Deviations

- In sample of 1000 people blood types (M,N,MN)
  - Genotype Observed Expected
    - MM 298 ?
    - MN 489 ?
    - NN 213 ?
- Estimates
  - \( p = \frac{(2\times298) + 489}{2000} = 0.5424 \)
  - \( q = \frac{(2\times213) + 489}{2000} = 0.4575 \)
  - \( p^2 = 0.2943 \)
  - \( 2pq = 0.4964 \)
  - \( q^2 = 0.2093 \)
Detecting Deviations

- In sample of 1000 people blood types (M,N,MN)
  - Genotype  | Observed | Expected
  - MM        | 298      | 294
  - MN        | 489      | 496
  - NN        | 213      | 209
- Compute Chi-Square
  - $\sum((\text{Observed}-\text{Expected})^2/\text{Expected} = .22$
  - Test against $p = .05$ value (3.84)
  - Conclude in equilibrium
- **Most important part - compare observed and expected frequencies of genotypes

Sources of Disequilibrium

- Mutation
  - Point mutations are rare
  - Usual in “neutral” locations
  - Other forms may be more common
  - Can be a factor in rapidly reproducing simple organisms (viruses)
Sources of Disequilibrium

• Non-random mating
  – Sexual selection - choosy females
  – Assortative mating - like types
  – Inbreeding

Sources of Disequilibrium

• Genetic Drift
Sources of Disequilibrium

- Immigration and Emigration

Sources of Disequilibrium

- Selection
  - Directional
  - Disruptive
  - Stabilizing
Selection and Evolution

Source of Polymorphism (Genetic Diversity)

- Hidden Recessive Alleles
  - Many rare genetic diseases
- Heterozygous Advantage
  - Sickle cell (in Africa)
- Frequency Dependent Selection
  - Host-parasite interactions (Red Queen)
- Neutral Variation
  - Hard to demonstrate
Case Study - Evolution and HIV/AIDS

• HIV enters white blood cells using two receptors
  – CD4
  – CCR5

• A mutation exists that causes CCR5 receptors to not exist = CCR5∆32

• ∆32 is recessive so must be ∆32/∆32 homozygous for protection

Will ∆32 Save Africa?

• If current frequencies of infection and allele (.2) are high and 100% of ∆32 survive, while 75% of non ∆32 survive

• But this is not what is observed…
Will $\Delta 32$ Save Africa?

- Distribution of allele varies geographically

- In Western Europe frequency of allele is high but disease is low

\[
\begin{array}{c|c|c|c}
\text{Allele} & \text{Initial frequency} & \text{Fraction surviving} \\
\hline
+/+ & 0.2 & 0.995 \\
+/\Delta 32 & 0.995 \\
\Delta 32/\Delta 32 & 1.0 \\
\end{array}
\]

- So there isn’t much selective pressure to spread
Will $\Delta 32$ Save Africa?

- In Sub-Saharan Africa frequency of allele is low but disease is high
  
  \[ \begin{array}{cccc}
  \text{Generation} & 0 & 10 & 20 & 30 & 40 \\
  \text{Frequency of} & 0.2 & 0.6 & 1.0 & 0.75 & 0.75 & 1.0 \\
  \text{CCR5-$\Delta 32$ allele} & \\
  \end{array} \]

- So most copies of allele are in heterozygotes and there is little for selection to act upon

Selection within Individuals

- Individuals can be host to populations (as in HIV)
- AZT resistance in HIV (within individuals)
Selection within Individuals

- Even without drugs, HIV competes within its host
- Should more “virulent” strains evolve?
- Virulence
  - Ability to cause disease
- Increase in virulence = spread faster within the host

Selection on HIV

- Increase in virulence = spread faster between hosts?
  - Perhaps
- So should HIV increase/decrease in virulence?
- What about transmission between individuals?
  - Need to understand disease dynamics...
Epidemics - SIR Models

- Susceptible – Infected – Recovered (Anderson and May)
- \( \frac{dS}{dt} = -\beta SI \)
- \( \frac{dI}{dt} = \beta SI - \gamma I \)
- \( \frac{dR}{dt} = \gamma I - \beta I \)
  - \( \beta \) = transmission rate
  - \( \gamma \) = recovery rate
- \( R_0 = \frac{\beta}{\gamma} \)
  - Average number of individuals that each sick individual infects

\[ SIR \text{ Disease Dynamics} \]

\( R_0 \) Predicts Spread

- \( R_0 > 1 \) results in spread
- The value of \( R_0 \) for some well-known diseases
  - Disease \( R_0 \)
  - AIDS 2 to 5
  - Smallpox 3 to 5
  - Measles 16 to 18
  - Malaria \( \gg 100 \)
SIR – High Mixing

- High movement rates for all individuals
- Sick people mix evenly with susceptibles and non-susceptibles
- Approximates differential equations and in this case everyone gets sick

SIR – Low Mixing

- Low movement rates for all individuals
- Sick tend to meet recovered and sick
- Leads to much lower infection rate and in this case only 50% total infection
- In many diseases (HIV) infected individuals are not as likely to meet uninfected individuals
Ever Increasing or Decreasing Virulence?

- Myxoma
- Introduced to control rabbit problem in Australia
- Original strains killed 95%+ of infected rabbits
- Later recovered predominant strains only killed 50%
- Sometimes increase in virulence decreases $R_0$

More Complete Definition of $R_0$

- $R_0 = \frac{\beta N}{\alpha + b + \gamma}$

$\beta =$ rate constant of infectious transfer (transmissibility)
$N =$ density of the susceptible host population
$\alpha =$ rate of parasite-induced mortality (virulence)
$b =$ rate of parasite-independent mortality
$\gamma =$ rate of recovery
Constraints on Virulence

- Virulence ($\alpha$) proportional to transmissibility ($\beta$)
- So intermediate values can be most successful
- Perhaps suggests that HIV will also remain at similar virulence, as it lets individuals pass it on.

Units of Selection

- Is this evidence of selection happening on groups (the populations within the hosts)?
- Are some viruses being “altruistic” and limiting the virulence?
- This isn’t necessary to explain. We can think about it from the “selfish” point of view.