HIV and population genetics

2

Please sit in row H or forward
Random biological fact of the day: a giant wall of... myovirus P-SSM4

>10^8 viruses / mL in seawater

Topics for today

• HIV evolution inside a patient (the four forces continued)
• Measures of genetic variation in a population: heterozygosity
Where we left off last time...

- Mutation
- **Genetic drift**
- Natural selection
- Migration
A simple model of allele frequency changes due to genetic drift

- Two alleles
- Fixed population size
- Organisms have only one copy of each sequence
- Asexual reproduction
- Make next generation by sampling with replacement
- No new mutation
Repeatedly sample until next generation is same size as first.
Genetic Drift

- Frequency of A1 allele
- A1 allele fixed
- Generations
- A1 allele lost
- Popsize 100
In a population of size 50, the A1 allele is present in a single copy. What is the probability that A1 is lost in the next generation?
Relating the model to the real scenario of HIV in a patient

Key similarity: random effects exert strong influence over how much a given viral particle is able to reproduce.
Four forces that can change allele frequencies in populations

• Mutation
• Genetic drift
• **Natural selection**
• Migration
HIV must struggle against the patient’s immune system
A definition of selection:

Natural selection is the differential survival and/or reproduction of individuals in a population due to differences in their physical characteristics.
A population genetic model with natural selection

Fitness:
- Metric of an individual organism’s ability to survive and reproduce
- In our model, the probability that an individual survives to reproduction

\[ w_1 = 1 \quad w_2 = 0.5 \]
Natural selection
Natural selection
In addition to natural selection, which of the four forces is present in this model?
Natural selection: alleles which lead to better survival become more numerous over time.
Natural selection within an HIV patient: measuring relative fitness

- Allow strains to compete with each other over time.
- Measure amount of each at end:

Cell culture dish with helper T cells from a person without HIV

Strain 1

Strain 2

Strain 1

Strain 2

Natural selection within an HIV patient: relative fitness over time
Four forces that can change allele frequencies in populations

• Mutation
• Genetic drift
• Natural selection
• Migration
Migration (gene flow)

Migration can
• Introduce new alleles
• Alter freq of existing alleles
Topics for today

• HIV evolution inside a patient (the four forces continued)
• Measures of genetic variation in a population: heterozygosity
Population genetic simulations with ongoing mutation

- Sampling with replacement
- Replication with possible mutation
Measuring genetic variation in a population: heterozygosity

Heterozygosity: the probability two alleles sampled randomly (with replacement) from the gene pool are different.
Measuring genetic variation in a population: heterozygosity

Heterozygosity: the probability two alleles sampled randomly (with replacement) from the gene pool are different.
Calculating heterozygosity

<table>
<thead>
<tr>
<th>Alleles at site 3</th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Probability of pulling 2 A alleles</th>
<th>Probability of pulling 2 C alleles</th>
<th>Probability of pulling 2 G alleles</th>
<th>Probability of pulling 2 T alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterozygosity at site 3 (probability of pulling two different alleles)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you have a little extra time: what is the theoretically maximum possible value for heterozygosity in nucleotide sequences?
Calculating heterozygosity

<table>
<thead>
<tr>
<th>Alleles at site 3</th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>frequency</td>
<td>0.2</td>
<td>0.4</td>
<td>0.4</td>
<td>0</td>
</tr>
</tbody>
</table>

Heterozygosity at site 3 (probability of pulling two different alleles)

1 - 0.04 - 0.16 - 0.16 - 0 = 0.64

If you have a little extra time: what is the theoretically maximum possible value for heterozygosity in nucleotide sequences?

(1 - .25^2 - .25^2 - .25^2 - .25^2) = 0.75
Measuring genetic variation in a population: heterozygosity

<table>
<thead>
<tr>
<th>Heterozygosity, site 1</th>
<th>Heterozygosity, site 2</th>
<th>Heterozygosity, site 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Average heterozygosity over all three sites: 0.213
Plot of heterozygosity in a drift simulation with ongoing mutation

- Starting allele: AAA
- Mutation probability: 0.0033 mutations per nucleotide
- Heterozygosity averaged over three nucleotide sites

Assymptote less that 0.75
Mutation probability = 0.0033 mutations per nucleotide

1. Why does this plot of heterozygosity vs. generations have an asymptote less than 0.75?

2. What would you expect to happen if we increased the mutation probability? Decreased it? (you can draw this)
1. Why does this plot of heterozygosity vs. generations have an asymptote less than 0.75?

   Genetic drift is removing variation in these simulations. The more variation there is, the more drift removes. Once the amount of variation becomes large enough, the variation added by mutation is balanced by that removed by drift.

2. What would you expect to happen if we increased the mutation probability? Decreased it?
   (you can draw this)
Question: why has HIV killed so many people?

- Defeating the immune system
  - Attacks the immune system itself
  - Rapid evolution

Hand in your worksheets please!

(and be sure you put your name on it)