Review

Genetic drift

Evolutionary force that removes genetic variation from a population
Strength is inversely proportional to the effective size of the population
If genetic variation is continuously removed, why does variation still exist?
Mutation

DNA can mutate and recombine (premise 2)

Mutation: error in replication of DNA across generations

  ultimate source of genetic variation

Prob. of losing a newly arisen mutation:

\[
\sum_{i=0}^{2} \frac{e^{2i}}{i!} \cdot \frac{1}{2^i} \cdot e^{\frac{1}{i!}} \cdot e^{1} = 0.37
\]

Regardless of population size, over a third of newly arisen mutations are lost in one generation due to drift.

Genetic variation

Mutation and genetic drift interact to shape the genetic variation of a population

How much variation can be found in natural populations?

What is the significance of that variation?

How exactly do mutation and genetic drift interact?
How much variation?

Two schools of thought both arose from the *Drosophila* laboratory of Thomas Hunt Morgan.

Two schools

Argued over the amount of variation in natural populations

The *classical school*: little variation

- Thomas Hunt Morgan
- H. J. Muller

The *balanced school*: lots of variation

- Alfred Sturtevant
- Theodosius Dobzhansky

Affected by the limited techniques at the time
Classical school

Morgan and Muller scored variation by inbreeding *Drosophila* to reveal single-locus variants with visible morphological effects.

Variation was required to even identify the locus at all.

Saw that most individuals were alike for those traits.
Common homozygous “wild type” vs rare heterozygous “mutant” allele.
Deleterious mutants are quickly lost, beneficial ones go to fixation.
Selection limits the amount of genetic variation in a population.

Balanced school

Sturtevant and Dobzhansky scored variation by staining giant polytene chromosomes of the larval salivary glands.
They looked at genetic variation in natural *Drosophila* populations, not inbred lab strains.

Chromosomal inversions inherited as Mendelian traits, just like Morgan and Muller's morphological traits.
Balanced school

Found chromosomal inversions inherited as Mendelian traits
Found extensive diversity in natural populations (no wild type)
Proposed that diversity is maintained by balancing selection

Two schools

Both schools were limited by the techniques available to score genetic variation

Looked at different types of variation
Required variation to exist for loci to be detected
Specific genes couldn't be scored directly
  Couldn't determine how many were variable and how many weren't

Both focused on selection as the primary mechanism governing the amount of variation
Two new techniques

Two new types of data became available in the 1960s

- Protein electrophoresis
- Amino acid sequencing

Genetic variation could be scored at the level of proteins / gene products

A new field of molecular evolution arose

Protein electrophoresis

Non-denatured proteins are placed in a polyacrylamide gel, which is subjected to an electrical current.

Proteins with different sizes or charges move at different rates.

Variation for a specific gene product could be scored directly across individuals in a population.
Protein electrophoresis

Lewontin & Hubby (1966), Johnson et al. (1966), and Harris (1966) showed that about 1/3 of all protein coding loci were polymorphic for electrophoretically detectable alleles in Drosophila and in humans.

Far more genetic variation than initially expected!
Not consistent with the classical school

### Table 5

<table>
<thead>
<tr>
<th>Population</th>
<th>No. of loci polymorphic</th>
<th>Proportion of loci polymorphic</th>
<th>Proportion of genome heterozygous per individual</th>
<th>Maximum proportion of genome heterozygous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strawberry Canyon</td>
<td>6</td>
<td>.33</td>
<td>.148</td>
<td>.173</td>
</tr>
<tr>
<td>Wildrose</td>
<td>5</td>
<td>.28</td>
<td>.106</td>
<td>.156</td>
</tr>
<tr>
<td>Gimmecon</td>
<td>5</td>
<td>.28</td>
<td>.099</td>
<td>.153</td>
</tr>
<tr>
<td>Mather</td>
<td>6</td>
<td>.33</td>
<td>.148</td>
<td>.173</td>
</tr>
<tr>
<td>Flagstaff</td>
<td>5</td>
<td>.28</td>
<td>.081</td>
<td>.120</td>
</tr>
</tbody>
</table>

Average

.30
.115
.155

Amino acid sequencing

Homologous protein sequences could be determined directly

Homology: traits derived from a common ancestor
Variation *between* species could be measured

<table>
<thead>
<tr>
<th>Divergence in 1 chain of hemoglobin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse</td>
</tr>
<tr>
<td>Human</td>
</tr>
<tr>
<td>Mouse</td>
</tr>
<tr>
<td>Chicken</td>
</tr>
<tr>
<td>Newt</td>
</tr>
<tr>
<td>Carp</td>
</tr>
</tbody>
</table>

Note: Pairwise difference between humans and sharks nearly the same as between carps and sharks. Why?
Amino acid sequencing

Homologous protein sequences could be determined directly

Homology: traits derived from a common ancestor

Variation *between* species could be measured

<table>
<thead>
<tr>
<th></th>
<th>Mouse</th>
<th>Chicken</th>
<th>Newt</th>
<th>Carp</th>
<th>Shark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>16</td>
<td>35</td>
<td>62</td>
<td>68</td>
<td>79</td>
</tr>
<tr>
<td>Mouse</td>
<td>39</td>
<td>63</td>
<td>68</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Chicken</td>
<td>63</td>
<td>72</td>
<td>83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newt</td>
<td>74</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carp</td>
<td></td>
<td>85</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sequence divergence is approximately proportional to the time of divergence, not to divergence in phenotype.

Molecular changes occurred at a more or less constant rate: the molecular clock (Zuckerkandl and Pauling, 1962)

Also inconsistent with the classical school, which claimed that molecular changes are driven by adaptation (which should change over time)
Significance of variation?

The classical school was no longer tenable

The debate shifted from the amount of genetic variation, to the significance of genetic variation

The classical school became the school of neutral evolution

   Leading proponent and developer was Motoo Kimura (1968)

   Explored more fully by Jack King and Thomas Jukes (1969)

The neutral theory

The classical theory allowed two types of mutations:

1. Rarely occurring deleterious mutations
2. Extremely rare beneficial mutations

Kimura's neutral theory added a third mutational class to the classical school model:

3. Neutral mutations that have no impact on fitness
   
   The single wild type allele was replaced by a set of functionally equivalent alleles

   Note: The neutral model does NOT assume all mutations are neutral, and accepts the classical model for non-neutral mutations
Effect of 50 spontaneous mutation lines derived from a strain of yeast grown in a laboratory environment (Zeyl & DeVisser 2001)

The neutral theory

Explains the higher-than-expected level of genetic variation across phenotypically homogenous populations

Explains the molecular clock

Genetic variation was no longer governed solely by selection, but by mutation and drift

How?
Interaction of mutation and drift

Genetic drift fixes one allelic type and removes all others:

Probability of fixation of a new allele under neutrality \(\frac{1}{2N}\)

Mutation introduces new alleles into a population:

Rate of introduction of new alleles \(2N\)

Assume the infinite-alleles model: all mutations yield a new allele
\(\mu\) is the probability of a mutation occurring per generation per locus

Rate of neutral evolution (the substitution rate):

\[
\frac{1}{2N} \quad 2N
\]

Interaction of mutation and drift

\[
\frac{1}{2N} \quad 2N
\]

Large populations have a smaller probability of fixation, but introduce mutants at a greater rate.

Population size has no effect on the rate of evolution!

Drift is an important evolutionary force for neutral alleles in all populations, not just small ones!
The neutral theory

Explains the molecular clock

Rate of change $= \mu$

Governed by an internal property of the system, rather than an external property affected by changing environments.

The neutral theory

Explains the high level of genetic variation:

$$F \cdot t \quad \frac{1}{2N} \quad 1 \quad \frac{1}{2N} \quad F \cdot t \quad 1 \quad 1 \quad 2$$

- Probability of identity by descent at generation $t$
- Probability of identity by descent due to genetic drift
- Probability of no new mutation in both gametes
The neutral theory

Explains the high level of genetic variation:

At equilibrium: \( F_t \ F_t \ 1 \)

\[
F_{eq} = \frac{1}{2N} \left( \frac{1}{1^2} \ 1 \ 1 \right) \frac{1}{4N} \left( \frac{1}{1} \right)
\]

Let: \( 4N_{ef} \) so: \( F_{eq} \frac{1}{1} \)

Expected heterozygosity: \( H_{eq} \ 1 \ F_{eq} \frac{1}{1} \)

The neutral theory

Explains the high level of genetic variation:

Depending on the exact value of \( \theta \), the neutral theory can explain any degree of genetic variability found in a population.

\[
H_{eq} \frac{1}{1} \]

\( \theta = 4N_{e} \mu \)
The neutral theory

Explains the high level of genetic variation:

1/2N mutations go to fixation and transiently contribute to genetic variation

Most mutations are lost and contribute little to polymorphism levels

Time period of transient polymorphism

Critique of the neutral theory

Most observations below this threshold

Implies a small range of population sizes, and that almost all species have $N < 5,000$ (including insects and bacteria!)
Critique of the neutral theory

Tomoko Ohta developed the nearly neutral theory to explain this inconsistency. She showed that slightly deleterious mutations are effectively neutral in small populations, due to strength of genetic drift. With larger population sizes, genetic drift has less influence, and selection dominates, slowing the substitution rate. Larger population sizes therefore depress the amount of observable genetic variation. Unfortunately, this undermines the prediction of a strict molecular clock.

Significance of the neutral theory

Helps explain observed patterns of intraspecific variation and interspecific divergence. Shows that genetic drift plays an important role in large and small populations. Challenged the emphasis on Darwinian evolution. Now commonly used as a null model for testing the hypothesis of natural selection.