Neutral theory of molecular evolution

• Kimura 1968; King and Jukes 1969
• Claimed:
  • Most new mutations are deleterious and are lost immediately
  • Most of the observed molecular polymorphism and substitutions are neutral

Claimed that this is consistent with:
- High levels of genetic polymorphism
- The molecular clock
The rate of genetic drift is higher in smaller populations. Allele frequencies do drift in very large populations but at a very slow rate.
Loss of heterozygosity

• In the absence of new mutations, alleles drift to either loss or fixation. Thus the amount of heterozygosity within the population decreases a rate inversely proportional to population size

Heterozygosity = Fraction of sites that are heterozygous
Loss of heterozygosity
Loss of heterozygosity

- In the absence of new mutations, alleles drift to either loss or fixation. Thus the amount of heterozygosity within the population decreases at a rate inversely proportional to population size

\[
H_{t} = H_{0} e^{-2Nt}
\]

Heterozygosity = Fraction of sites that are heterozygous
• Profound effects of short-term reduction in population size.
• The effect of population bottlenecks (founder effects) on polymorphism

Rare alleles likely to be lost
But some rare alleles can by chance drift to high frequency within the population.
Bottleneck in Northern Elephant Seals

Figures from Bergstrom and Dugatkin. Evolution. 2nd Ed

mtDNA diversity over time

- Northern elephant seal
- Southern elephant seal

Observed heterozygosity $H_0$

Fraction of loci showing polymorphism

<table>
<thead>
<tr>
<th>Number of Individuals</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>Other types</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000–1800</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1892</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1914–1980</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1990–2000</td>
<td>107</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
During bottlenecks rare alleles can by chance reach high frequency

5% of islanders on Pingelap have A recessive form of achromatopsia (total colour blindness)
The allele is at ~20% frequency

Higher incidence of polydactyl in Amish population

Using Genetic Tests, Ashkenazi Jews Vanquish a Disease

A number of years ago, five families in Brooklyn who had had babies with a devastating disease decided to try what was then nearly unthinkable: to eliminate a terrible genetic disease from the planet.

The disease is Tay-Sachs, a progressive, relentless neurological disorder that afflicts mostly babies, leaving them mentally impaired, blind, deaf and unable to swallow. There is no treatment, and most children with the disease die by 5.

Tay-Sachs disease (a recessive disorder) is found at higher frequency in Ashkenazi Jews
Mutation-drift equilibrium (MDE)

Mutation-drift equilibrium (MDE)

- Variation lost by drift = variation introduced by mutation

- $N_e = \text{effective population size,}$
- $\mu = \text{mutation rate to new neutral alleles}$

At MDE, $\text{heterozygosity} = H \approx 4 \ N_e \mu$

The rate of genetic drift is slower in larger populations and the input of new mutations is higher. Large populations have a higher level of neutral polymorphism.
Neutral theory of molecular evolution

• Kimura 1968; King and Jukes 1969
• Claim: Most of the observed molecular polymorphism is neutral

Consistent with Neutral theory
Levels of genetic variability within Populations (i.e. heterozygosity) are high and are broadly correlated with population size
We can estimate effective population sizes $H=4N_e\mu$:

<table>
<thead>
<tr>
<th></th>
<th>$H$</th>
<th>$\mu$ (per gen)</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>0.1%</td>
<td>$2\times10^{-8}$</td>
<td>$10^4$</td>
</tr>
<tr>
<td>Chimpanzee</td>
<td>0.3%</td>
<td>$2\times10^{-8}$</td>
<td>$3\times10^4$</td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em></td>
<td>2%</td>
<td>$5\times10^{-9}$</td>
<td>$10^6$</td>
</tr>
</tbody>
</table>

Human census population~ $7\times10^9$; Chimp~ $2\times10^5$

Do these low effective population sizes reflect repeated bottlenecks

Or is neutral theory only partially correct?
Effective Population size

Real populations deviate from many of the assumptions of the Wright-Fisher model (even ignoring selection, population structure).

• The rate of genetic drift, i.e. loss of heterozygosity, will be increased if:
  • The population often crashes to a small size
  • Or if there is a high variance in reproductive success.

• The effective population size $N_e$
  • The size of an ideal population in which drift occurs at the same rate as that in an actual population

\[ N_e \ll N \]

In many (not all) situations, departures from model assumptions can be captured by substituting $N_e$ for $N$, i.e., by plugging in a fictitious $N$ that leads to the same level of genetic drift as observed.
Effective Population size

- If the population often crashes to a small size = rate of drift is higher
Effective Population size

- higher variance in reproductive success = rate of drift higher

Circled individuals have much higher reproductive success (independent of genotype at locus).
Coalescent process w. 2 lineages

Time back to coalescence (13 generations)