The Neutral Theory of Molecular Evolution

- Kimura (1968)
  - initially viewed as a challenge to Darwinian evolution
    - e.g., King & Jukes 1969 *Science* 164:788-798.
- many genetic polymorphisms have no effect on fitness and are therefore selectively neutral
- neutral polymorphisms are maintained by the combined effects of mutation and drift
  - mutations introduce new alleles as others are lost through drift

Origins of the "Selectionist-Neutralist" Debate

- the only "mutations" early biologists saw were ones that had phenotypic effects
- 1953 - structure of DNA (Watson & Crick / *Franklin*)
- 1960-70's - protein electrophoresis
  - revealed allelic diversity for many genes
- 1968 - Motoo Kimura - the neutral theory
  - motivated by allozyme (amino acid) variation
- discovery of "junk DNA"
  - 98.5% of human genome is non-coding
- DNA sequencing has revealed substantial silent and non-coding variation, suggesting that much genetic variation is selectively neutral (or nearly so!)

Infinite Alleles/Sites Model

- what is the expected level of genetic diversity (heterozygosity) given mutation and drift in a finite population?
- suppose a gene is 900 base pairs long, coding for 300 amino acids
  - there are $4^{900} = 10^{542}$ possible sequences (sorta…)
- thus, we can reasonably assume that each new mutation generates a unique allele…

Infinite Alleles/Sites Model

- it follows that alleles with the same sequence are identical by descent
- autozygous - a genotype with two alleles that are identical by descent
- allozygous - a genotype with alleles that are not identical by descent (is this possible?)
  - arbitrarily declare all alleles unique at $t = 0$
- autozygous = homozygous under the infinite alleles model
  - thus, the level of heterozygosity can be predicted from the expected level of autozygosity

Infinite Alleles/Sites Model

- F_t = probability that two randomly chosen alleles are *IBD*
  - same as autozygosity if we randomly choose alleles to form genotypes

$$F_t = \left(\frac{1}{2N}\right)(1-\mu)^2 + \left(1-\frac{1}{2N}\right)(1-\mu)^2 F_{t-1}$$

  - in this model, mutations generate new alleles and "erase" *IBD*