LECTURE 12

Selection (continued):

5. The Neo-Classical Perspective [T. Dobzhansky et al.]

   a) A corollary of the "classical" perspective was that there would be a "drive" towards homozygosity for the "best" allele at each locus and that heterozygotes were transient "intermediates"; this predicts that homozygotes should be common in natural populations and that frequencies of heterozygotes should decrease over time

   b) The "classical" perspective prevailed until the advent of biochemical techniques (protein electrophoresis, initially) that permitted assessment of more than a few genes; today, many population genetics laboratories can screen hundreds of genes/DNA sequences if time/money permits

   c) What has been found is that frequencies of heterozygotes in natural populations are in far higher than was expected, and moreover, that frequencies of heterozygotes appear to be fairly stable and do not necessarily decrease over time (at least in large populations)

   d) Some examples:

Taxon	Mean proportion of loci polymorphic per population
Drosophila	>50%
Marine invertebrates	≈60%
Fish	30%
Rodents	20%
Large mammals	≈25%
Outcrossing plants	>45%

   Note: An individual typically could be heterozygous at 5-15% of loci assayed!

   e) This led to the hypothesis of heterozygote superiority, *overdominant selection*

   Single-locus model of overdominant selection:

	AA	Aa	aa
	p^2	2pq	q^2
fitness	$1-S_A$	1	$1-S_a$
selection	$p^2-S_Ap^2$	2pq	$q^2-S_aq^2$

   S_A and S_a are selection coefficients, ranging from 0 to 1

f) Why should heterozygotes be at an advantage? There are several hypotheses and a large base of somewhat indirect data supporting one or more of the following:

Ecological theories (more balanced response to environmental shifts)
Variability theories (pre-adaptiveness)
Regulation theories (retains potential to alter regulation of pathways)
Biochemical theories (pathways can retain potential to flux)

g) Under an overdominant model of selection, an equilibrium eventually will be reached: it will be a balance point that maximizes frequencies of heterozygotes as a function of selection intensity *against* the two homozygous genotypes

$q = S_A / (S_A + S_a)$ $p = S_a / (S_A + S_a)$

h) Given the equilibrium, selection will be "directional" or "normalizing" (stabilizing) depending on how far the population is from equilibrium

i) The crucial point to be made about overdominant selection is that potentially there is a genetic disability suffered by the population -- called *segregational load*

at equilibrium, where $q = 0$: $(S_A)(S_a) / (S_A + S_a)$ = segregational load

(i) segregational load is defined as the proportion of genetic-lethal equivalents per generation and includes less fit individuals of both homozygous genotypes

(ii) A good example in humans is sickle cell anemia

S^NS^N = normal
S^NS^S = sickle cell "trait"
S^SS^S = sickle cell anemia

(iii) Selection normally would be against the recessive genotype; however, in regions where malaria is pandemic, selection favors the heterozygous genotype and is overdominant

(iv) In regions where there is no malaria, selection is against the homozygous recessive genotype (a "genetic" lethal); eventually, the population will come to the selection-mutation equilibrium, defined previously as…

$$q = (\mu/s)^{1/2}$$

(v) For purposes of example, assume $\mu = 10^{-6}$ and $s = 1.0$

$q = (10^{-6}/1)^{1/2}$ $q = 10^{-3}$ $q = 0.001$

$q^2 = 10^{-6}$ or 1 in a million individuals has sickle cell anemia

(vi) However, where malaria is prevalent, selection is against both homozygous genotypes, i.e., overdominant; in this case, the equilibrium value of q (i.e., when q = 0) will be…..

$q = S_A / (S_A + S_a)$

Assume: $S_A = 0.5$ (1/2 p^2 homozygotes die), and
$S_a = 1.0$ (it's still a recessive lethal in a genetic sense)

$q = 0.5/1.5$ $q = 0.33$ $q^2 = \approx 0.1$ (1 in 10 individuals)

Note: By eliminating 50% of the "normal" homozygotes (S^NS^N), the frequency of a *recessive lethal* increased from 0.001 to 0.330 (over 300-fold) and the incidence of sickle cell anemics in the population increased from 1 in a million to one in ten (that's six orders of magnitude)

Assume: $S_A = 0.25$ (1/4 p^2 homozygotes die), and
$S_a = 1.0$

$q = 0.25/1.25$ $q = 0.20$ $q^2 = 0.04$ ($\approx$ 1 in 20 individuals)

6. Genetic Load: Segregational load + mutational load + substitutional load

a) The reason why populations/individuals never really attain maximal fitness and why deleterious recessives (e.g., recessive lethals) remain in genomes

7. The Neutralist Hypothesis (M. Kimura and others)

a) The concept of segregational load implied a tremendous "load" on populations if one considered the proportion of polymorphic loci and assumed that the polymorphism was maintained by overdominant selection

(i) The "load" would be compounded locus-by-locus, and considering organisms (e.g., large mammals) where the number of loci is in the range of 30,000-40,000 genes (with 20-25% polymorphic), the question was "how was anyone left alive?"

b) The neutralist hypothesis was simply that allelic variants detected at many (most?) of the loci typically surveyed biochemically were effectively equivalent in terms of selection (i.e., "neutral" to natural selection in terms of functional equivalency)

c) In addition, a number of "mutations" should exist (e.g., at third positions of many codons) that would be expected to produce proteins with the same amino acid sequence; this is referred to as *degeneracy* of the genetic code

d) Perhaps equally important was the notion that if these variants were neutral, their frequencies in populations would essentially be a function of mutation rate, effective population size, and *time*; this (ultimately) led to the hypothesis of a "*molecular clock*" (still a very controversial issue)

8. Frequency-dependent selection

a) Another model proposed to explain observed levels of heterozygosity in natural populations

b) The concept is that the frequency of phenotypes (genotypes) essentially determines fitness values, largely as a function of resource availability. A good example in the textbook involves "mimicry" in butterflies. Another example involves alcohol dehydrogenase (ADH) variants in populations of *Drosophila* living near wineries.

ADH^f = metabolizes long-chain alcohols
ADH^s = metabolizes short-chain alcohol

Frequency of either allele fluctuates with utilization/availability of different alcohols

c) A third example is *industrial melanism* in the modern times, where both light and dark "alleles are maintained in balance in the population as a whole (i.e., higher frequency of homozygous "darks" near industrial centers, but higher frequencies of "lights" in the country-side.