Chapter 26 Lecture Notes: Population Genetics

I.   Introduction
      A. Darwin's Theory of evolution
            1. Variation: Among individuals in a population there is phenotypic and
            genotypic variation
            2. Heredity: Offspring are more similar to their parents than to unrelated
            individuals
            3. Selection: Individuals having some phenotypes are more successful at
            surviving and reproducing than others
      B. Population genetics
            1. the translation of Darwin's three principles into genetic terms
            2. study of heredity of traits controlled by one or a few genes in a group of
            individuals
                  a) description of genetic structure of a population (patterns of genetic
                  variation found among individuals in a group)
                  b) examination of how genetic structure varies in space and time
                  c) evaluation of the processes that are responsible for producing genetic
                  variation
            3. practical applications include importance for conservation biology and
            biodiversity
            4. good example of the uses of mathematical theory in biology

II.   Genetic Structure of Populations
      A. Genotypic frequencies

            f (particular genotype) = # individuals with that genotype / total # individuals

      B. Allelic frequencies
            1. Frequently used over genotypic frequencies in calculations because
                  a) Genotypes break down to alleles when gametes are formed
                  b) Alleles (not genotypes) are passed on to progeny
                  c) Fewer alleles than genotypes so there are less parameters in the
                  equations
            2. $f(A) = p$ and $f(a) = q$ AND $f(A) + f(a) = p + q = 1$
            3. Calculation from observed # of individuals with each genotype (Using A and
            a as alleles of 1 gene)

            f (allele) = # of copies of the allele in a population / sum of all alleles

            $$p = f(A) = \frac{2 \ (\# \ AA \ individuals) + (\# \ of \ Aa \ individuals)}{2 \ (total \ \# \ of \ individuals)}$$

4. Calculation from the <u>frequencies</u> of the genotypes (Using A and a as alleles of 1 gene)

$$p = f(A) = f(AA) + \tfrac{1}{2}\, f(Aa)$$

5. Extensions of allelic frequency calculations: Multiple alleles (A1, A2, A3)

$$f(A1) = f(A1A1) + \tfrac{1}{2}\, f(A1A2) + \tfrac{1}{2}\, f(A1A3)$$

III. Genetic variation within a population
<u>Polymorphism</u> = genetic variation; the occurrence of several phenotypic forms of a character associated with one locus (gene) or homologues of one chromosome
   A. Types of genetic variation that population geneticists examine
      1. Morphological polymorphisms
      2. Chromosomal polymorphisms
      3. Immunological polymorphisms
      4. Protein polymorphisms
      5. DNA sequence polymorphisms
   B. <u>Heterozygosity</u> = measure of the frequency of the heterozygote genotype at a loci or at multiple loci

IV. The Hardy-Weinberg Equilibrium ($p^2 + 2pq + q^2$ = f(AA) + f(Aa) + f(aa) = 1)
   A. Simple explanation for how Mendelian principles that result from meiosis and sexual reproduction influence allelic and genotypic frequencies
   B. Assumptions
      1. Infinitely large population (where very large can be considered infinite) Smaller populations allow for chance deviations from the expected ratios which change allelic frequencies (see genetic drift)
      2. Randomly mating population for the trait that is being examined
      3. Population must be free from addition or subtraction of alleles due to:
         a) Mutation
         b) Migration
         c) Natural selection

C. Derivation:
For 2 alleles (A and a) of one gene, let $p = f(A)$ and $q = f(a)$.

|		eggs	
		$p = f(A)$	$q = f(a)$
Sperm	$p = f(A)$	$p^2 = f(AA)$	$pq = f(\text{``aA''})$
↓	$q = f(a)$	$pq = f(\text{``Aa''})$	$q^2 = f(aa)$

$$p^2 + 2pq + q^2 = f(AA) + f(Aa) + f(aa) = 1$$

D. Predictions of Hardy-Weinberg equilibrium
   1. The genotypic frequency is $p^2 + 2pq + q^2 = 1$ after 1 generation of mating.
   2. The frequency of alleles does not change over time.
   3. Rare alleles are virtually never in the homozygous condition.
E. Algebraic proof of H-W Equilibrium on reserve at the library under Hardy-Weinberg proof if you are interested
F. General relationship between homozygous and heterozygous frequencies and the frequency of each allele:
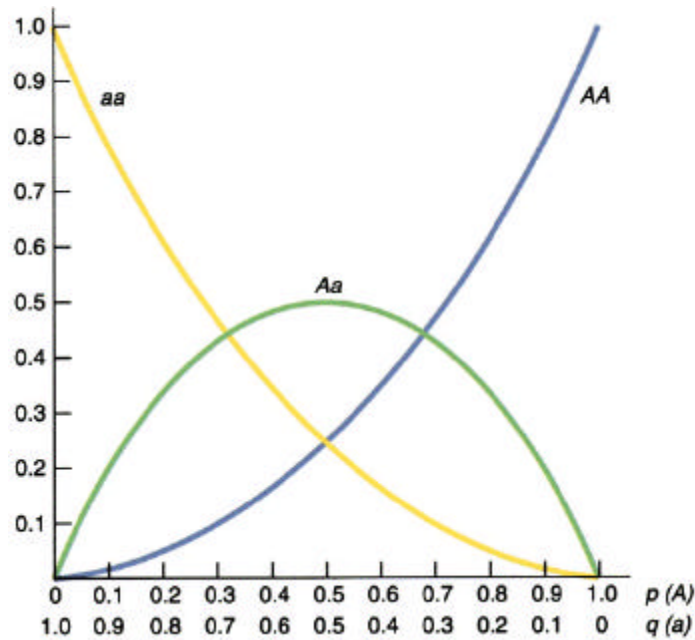
(From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)

G. Extension of the Hardy-Weinberg equilibrium: Multiple alleles
   For 2 alleles: $p^2 + 2pq + q^2 = (p + q)^2$
   For 3 alleles: $p^2 + 2pq + q^2 + 2pr + 2qr + r^2 = (p + q + r)^2$

H. Testing for Hardy-Weinberg equilibrium:

Consider the example where f(AA) = 0; f(Aa) = 1; f(aa) = 0

1. Calculate the observed allelic frequencies:

f(A) =   p   = f(AA) + ½ f(Aa)
                = 0 + ½ (1)
                = 0.5
f(a) =   q   = 1- p
                = 1 – 0.5
                = 0.5

2. Compute the expected genotypic frequencies based on H-W:

$f(AA) = p^2 = (0.5)^2 = 0.25$
$f(Aa) = 2pq = 2(0.5)(0.5) = 0.5$
$f(aa) = q^2 = (0.5)^2 = 0.25$

3. Compare expected genotypic frequencies from (2) with observed genotypic frequencies given. If expected = observed, then the population is at H-W equilibrium. If expected does NOT = observed, then the population is NOT at H-W equilibrium.

	expected	observed
f(AA)	0.25	0
f(Aa)	0.5	1
f(aa)	0.25	0

4. Thus, this population is definitely NOT at equilibrium. Sometimes if the numbers are closer together, it is necessary to do a chi-square analysis.

I. Using Hardy-Weinberg equation to estimate allelic and genotypic frequencies based on phenotypic frequencies.

Consider a population at Hardy-Weinberg equilibrium where the frequency of tallness (A-) is 0.96 and the frequency of shortness (aa) is 0.04.

1. Calculate the allelic frequency of the recessive allele:

$f(aa) = q^2 = 0.04$ → Thus, $q = 0.2$

2. Calculate the allelic frequency of the other allele:

p $= 1 - q$
$= 1 - 0.2$
$= 0.8$

3. Calculate the genotypic frequencies using the allelic frequencies from (2):

$f(AA)$ $= p^2$
$= (0.8)^2$
$= 0.64$

$f(Aa)$ $= 2pq$
$= 2(0.8)(0.2)$
$= 0.32$

$f(aa)$ $= q^2$
$= (0.2)^2$
$= 0.04$

V.   Sources of genetic variation
      Important because:
      • Determines the potential for evolution and adaptation
      • The ability for a population to persist over time may be influenced the amount of genetic variation it can draw upon in the event that the environment changes.

A. Mutation
The mutation rate influences the frequency of the alleles in the population. For the mutation from A → a

$$p_n \text{ approx.} = p_0 e^{-n\mu} \quad \text{(assuming } \mu \text{ is very small)}$$

p_n is f(A) after n generations p_0 is f(A) in the initial generation

n = # of generations μ = mutation rate from A → a

Derivation:
Consider the mutation from A to a, which occurs at a rate of μ.
The change in f(A) or p   = Δp:
Δp can be described by 2 equations: $p_t - p_{t-1}$
                                    and $-\mu \, p_{t-1}$
where t = a particular generation and t-1 = the previous generation
note that p decreases with each generation and so Δp will decrease with each generation (see graph below)*

$\Delta p = p_t - p_{t-1}$

$p_t = \Delta p + p_{t-1}$ Rearrangement of the formula

$p_t = -\mu \, p_{t-1} + p_{t-1}$ Substitution of $-\mu \, p_{t-1}$ for Δp

$p_t = p_{t-1} - \mu \, p_{t-1}$ Rearrangement of the formula

$p_t = p_{t-1}(1-\mu)$ Algebraic rearrangement of the formula

                          The value of p in a particular generation is the value of p in the previous
                          generation times $(1-\mu)$
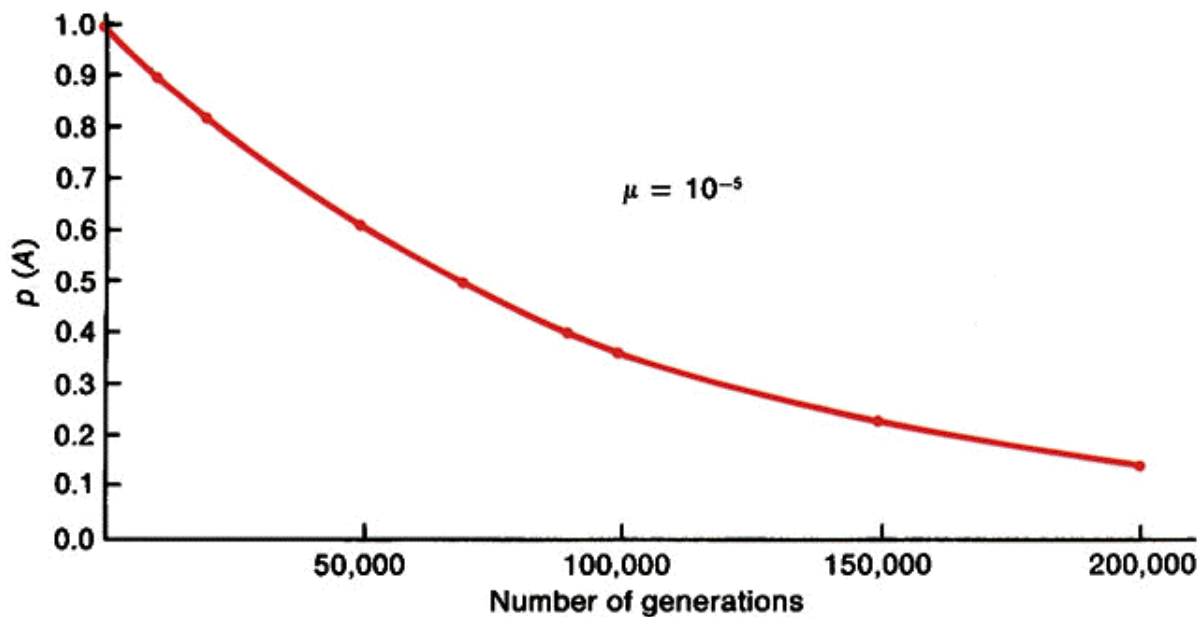
Now for the generation number 2   or t+1:
p_{t+1} $= (p_t)(1-\mu)$
           $= (p_{t-1}(1-\mu))(1-\mu)$ Substitution of $(p_{t-1}(1-\mu))$ for p_t
           $=(p_{t-1})(1-\mu)^2$ Rearrangement

$p_n = p_0(1-\mu)^n$ After many generations (n) where p_0 is the initial p

$p_n = p_0 e^{-n\mu}$ If μ is very small then $(1-\mu)^n \sim e^{-n\mu}$

At top of graph: $\mu = 10^{-5}$

Y-axis: $p\ (A)$ with values 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0

X-axis: Number of generations — 50,000, 100,000, 150,000, 200,000

(From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)

B. Recombination

C. Migration (individuals migrating into the population may introduce new alleles into the gene pool)

      1. Effects:

            a) Introduction of new alleles

            b) Changes in the allelic frequencies

            c) Counteracts genetic drift (see below)

      2. Calculations

Let p_t = f(allele in recipient population)

P = f(allele in donor population)

Let p_{t+1} = f(allele in next generation)

m = proportion of recipient population made up of new migrants

$$p_{t+1} = mP + (1-m)\,p_t \qquad \text{p from donor + p from recipient}$$
$$= mP + p_t - m\,p_t$$
$$= p_t + m(P - p_t)$$

$$\Delta p = p_{t+1} - p_t$$
$$= p_t + m(P - p_t) - p_t \quad \text{substitution of } p_t + m(P - p_t) \text{ for } p_{t+1}$$
$$= m(P - p_t) \qquad\qquad\quad \text{simplification}$$

D. Nonrandom mating
   1. Positive assortative mating – individuals with similar phenotypes mate preferentially (decreases heterozygosity)
   2. Negative assortative mating – individuals with different phenotypes mate preferentially
   3. Inbreeding – mating between related individuals occurs more frequently than predicted by chance (decreases heterozygosity)
      a) Measured in terms of the coefficient of inbreeding (F)
      b) Analysis using pedigrees
      c) In a closed population, founded by a small number of individuals, there will be a decrease in heterozygosity over time because some lineages (and genes) will die out. The more individuals there are in the population, the longer is takes for this to occur. The decrease in heterozygosity is counteracted by mutation and migration.
   4. Outbreeding– mating between related individuals occurs less frequently than predicted by chance

E. Genetic drift- random change in allele frequencies due to chance
   1. Causes:
      a) Small population size
      b) Founder effects – occurs when a population is initially established by small number of breeding individuals
      c) Bottleneck effect – occurs when a population is dramatically reduced in size
   2. Effects
      a) Decrease in allelic frequency over time because the f(allele) will either go to 0 or 1.
      b) Reduction in genetic variation

F. Natural selection
   1. <u>Adaptation</u> is the process by which traits evolve which makes the organisms more suited to its environments. Adaptation arises via <u>natural selection</u> which is the differential reproduction of certain genotypes because organisms carrying those certain genotypes are better suited to the environment.
   2. Natural selection is measured by assessing survival and reproduction rates which collectively is called fitness.
   3. W = Darwinian fitness = relative probability of survival and relative reproduction rate of a particular genotype or phenotype
   4. Calculation of how W affects allelic and genotypic frequencies:

      a) Preselection: $p^2 + 2pq + q^2 = f(AA) + f(Aa) + f(aa) = 1$
      b) During selection: Each genotype has a particular fitness W_{XX}
      c) Postselection:
      $f(AA)' = W_{AA}(p^2)$
      $f(Aa)' = W_{Aa}(2pq)$
      $f(aa)' = W_{aa}(q^2)$

d) Normalization to 1 because $W_{AA} + W_{Aa} + W_{aa} > 1$ by dividing each genotypic frequency by the total frequency of the population $\overline{W}$, where $\overline{W} = W_{AA}(p^2) + W_{Aa}(2pq) + W_{aa}(q^2)$

$f(AA)' = W_{AA}(p^2) / \overline{W}$

$f(Aa)' = W_{Aa}(2pq) / \overline{W}$

$f(aa)' = W_{aa}(q^2) / \overline{W}$

e) Calculation of postselection allelic frequencies:

$$p' = f(A)' \quad = W_{AA}(p^2)/\overline{W} + \tfrac{1}{2} W_{Aa}(2pq)/\overline{W}$$
$$= \frac{W_{AA}(p^2) + W_{Aa}(pq)}{\overline{W}}$$
$$= p\,\frac{[W_{AA}(p) + W_{Aa}(q)]}{\overline{W}}$$

$$q' = f(a)' \quad = W_{aa}(q^2)/\overline{W} + \tfrac{1}{2} W_{Aa}(2pq)/\overline{W}$$
$$= \frac{W_{aa}(q^2) + W_{Aa}(pq)}{\overline{W}}$$
$$= q\,\frac{[W_{aa}(q) + W_{Aa}(p)]}{\overline{W}}$$

f) Substitution of the fitness of the A allele (W_A) for $[W_{AA}(p)+W_{Aa}(q)]$

$p' = p\,(W_A) / \overline{W}$

$q' = q\,(W_a) / \overline{W}$

Note that after one generation, the $p' = p$ times the ratio of the fitness of the A allele to the mean fitness. Thus, if $W_A > \overline{W}$ then $(W_A / \overline{W}) > 1$ and p will increase. Conversely, if $W_A < \overline{W}$ then $(W_A / \overline{W}) < 1$ and p will decrease.

5. Relationships among the fitness values:
   a) $W_{AA} = W_{Aa} = W_{aa}$ → No selection
   b) $W_{AA} = W_{Aa} < 1$ and $W_{aa} = 1$ → Natural selection against a dominant allele
   c) $W_{aa} = W_{Aa} < 1$ and $W_{AA} = 1$ → Natural selection against a recessive allele

d) $W_{AA} < W_{Aa} < 1$ and $W_{aa} = 1$ → Natural selection is operating without effects of dominance

e) W_{AA} and $W_{aa} < 1$ and $W_{Aa} = 1$ → Natural selection is favoring the heterozygote

f) W_{AA} and $W_{aa} = 1$ and $W_{Aa} < 1$ → Natural selection is operating against the heterozygote

g) Important conclusions from a – f:
   (1) b, c, and d are directional selections which results in the elimination or great reduction in the frequency of one allele.
   (2) e does not result in evolutionary change once a stable equilibrium is reached.
   (3) f is very rare.
   (4) Natural selection against a completely recessive trait (where $W_{AA} = W_{Aa} = 1$ and $W_{aa} < 1$) proceeds very slowly because it can remain "hidden" in the heterozygote where selection will not act upon it. This is called <u>protective polymorphism</u>.

6. The rate of change of gene frequencies due to natural selection (Δp):

$$
\begin{aligned}
\Delta p \quad &= p' - p \\
&= [p\,(W_A / W)] - p & &\text{Substitution of } [p\,(W_A / W)] \text{ for p'} \\
&= p\,[(W_A / W) - 1] & &\text{Rearrangement} \\
&= p\,\frac{(W_A - W)}{W} & &\text{Rearrangement} \\
&= p\,\frac{[W_A - (pW_A + qW_a)]}{(pW_A + qW_a)} & &\text{Substitution of } pW_A + qW_a \text{ for } W_A \\
& & &\text{and some rearrangement} \\
&= pq\,\frac{(W_A - pW_a)}{W}
\end{aligned}
$$

This formula shows that Δp depends on the pq values and thus is proportional to the $f(Aa) = 2pq$. Thus, $>f(Aa)$ → $>\Delta p$ and $<f(Aa)$ → $<\Delta p$
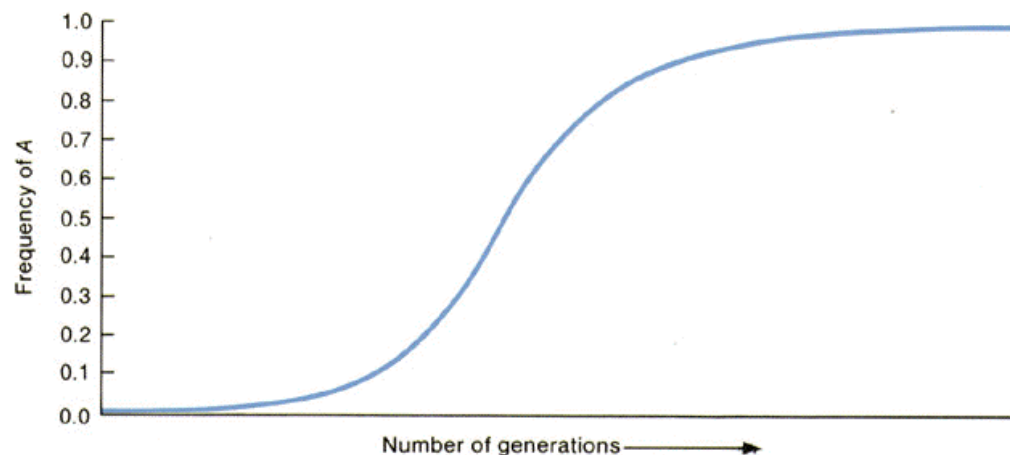
From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)

VI.   Balanced Polymorphism

$$\Delta p \quad = pq\frac{(W_A - pW_a)}{W}$$

If $W_A = W_a$ then $\Delta p = 0$. This can be due to the phenomenon of balanced polymorphism, whereby slight perturbations of p when $W_A = W_a$ results in the return to p. This will occur when the heterozygote has greater fitness than either homozygote.

See handout for derivation of the formula for p at equilibrium and for the example of the sickle cell anemia locus.

VII.   Artificial selection
A. Important in the laboratory, agriculture, and animal husbandry
B. Can be accomplished by truncation selection in which one pools individuals →
measures individuals → chooses individuals that meet a certain criterion → mates those
individuals above the cutoff → repeat many times
                a) Constant truncation (1 cutoff point used throughout
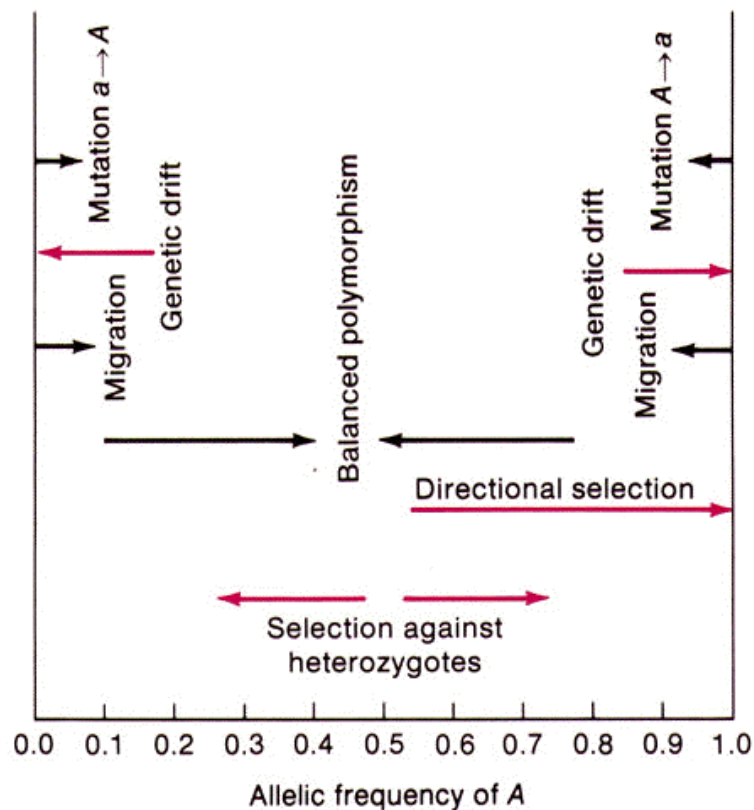   b) Proportional truncation (certain % used throughout)
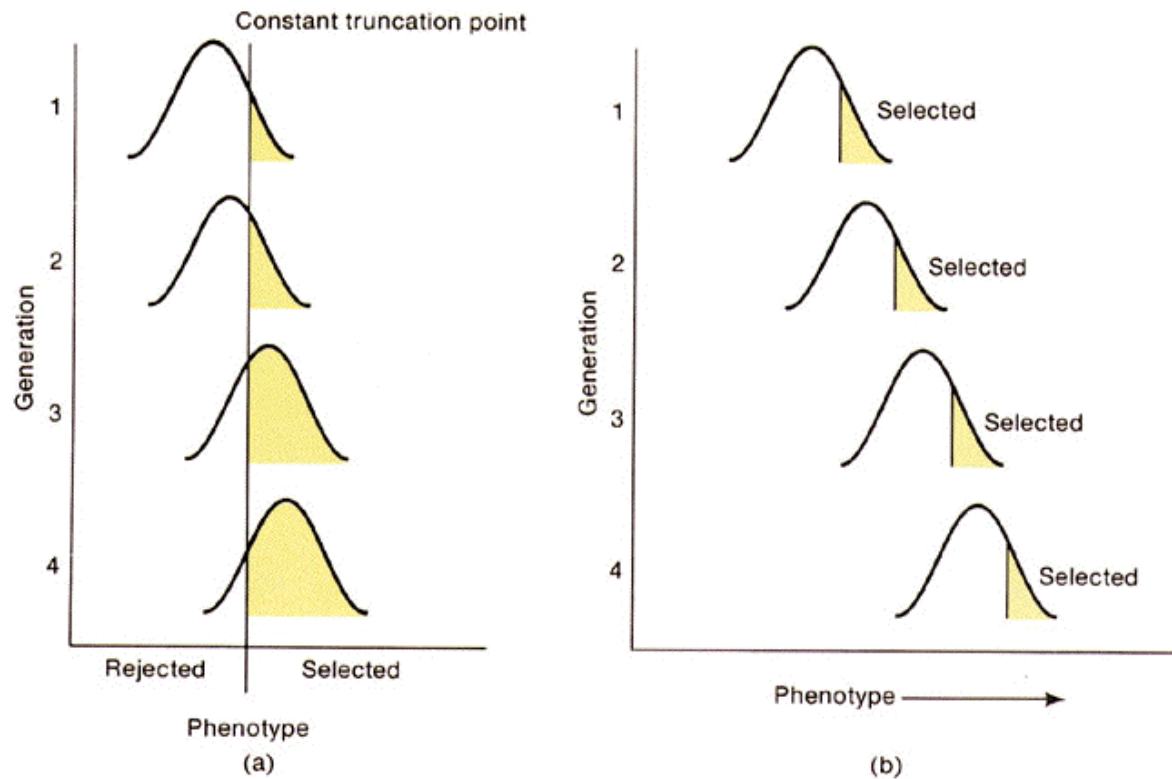
. From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki,
Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)

VIII.   A Synthesis of Forces:

Table 26-13   How the Forces of Evolution Increase (+) or Decrease (−) Variation within and between Populations

Force	Variation within populations	Variation between populations
Inbreeding or genetic drift	−	+
Mutation	+	−
Migration	+	−
Selection		
Directional	−	+/−
Balancing	+	−
Incompatible	−	+

Red arrows = decrease in variation   Black arrows = increase in variation

From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)