## Estimation of Mutation in Quantitative Traits:

- Mutation: the source of population genetic variation on which natural or aritifical selection acts (to produce genetic changes leading to adaptive evolution)
- Number of rounds of replicaiton from zygote to the formation of gametes varies between species and is why rates of mutations vary between species
- Men and women: the number of rounds of replication in ova stay at 22 while it continues to grow for spermatozoa in men
- However, chromosomal mutations increase in women because of cytoplasmic deterioration in oocytes over time
- Given the difficulty of identifying genotypes, it is only possible to obtain average estimates of the effect of mutations and their degree of dominance
- Mutational variance $V_M$ is the increase in additive genetic variance per generation due to mutation
- And thus mutational heritability is the ratio of $V_M/V_E$
- Probability of fixation of a mutation:
    - In finite populations, genetic drift may cause beneficial alleles to be lost and deleterious alleles to be fixed
    - If the mutational selection coefficient is small $(N_e s/N < 1)$ and $N_e s$ is high, then the probability of fixation can be approximated by the selective advantage s, so that most beneficial mutations of small effect are lost in a population of finite size
    - Neutral mutations have a $P_f$ equal to their initial frequency $1/2N$
    - Deleterious mutations are destined to be eliminated by selection
    - Mutations where absolute value of $N_e$ is substantially smaller than 1, the probability of fixation is the same for that of a neutral mutation. When mutation is advantageous or disadvantageous, the action of drift is more intense than selection and so the fate of the mutation is dependent on chance. This is referred to as quasi-neutral mutations.
    - Mutations can be beneficial, neutral or deleterious based on effective population size
- Estimating the rate of mutation and mutational effects
    - Balanced chromosome technique: cross an individual with another with a balanced chromosome (carrying inversions that inhibit recombinations along with visible markers for identification)
    - Cross wt male with If female to obtain copies of chromosome that are used to create multiple independent lines
    - Mutations accumulate in each line and disappear when deleterious heterozygous effects are sufficiently large
    - This approach has been applied to C. elegans, where lines are started by either self-fertilization or brother sister mating. Mutations accumulate in the genome, and because $N_e s << 1$, mutations behave as neutral and have a highly probability of fixation.
    - Mutations can be studied if they affect traits that are fitness components, either directly measured (productivity) or it may be in competition with the starting population (founder individuals in C. elegans can be frozen then thawed and compared with individuals from a line)
    - Accumulating mutations is expected to reduce the mean value of the trait will increasing the variance in that trait, and $\delta V/\delta M = \bar{s}(1+C^2)$, where U is the haploid genome mutation rate, and C is the coefficient of variation of mutational effects, and s is the average of selective effects of mutations weighted by their mutation rate (Bateman-Mukai method)

- Bateman-Mukai enables the estimation of mutational variance (increase in additive genetic variance per generation due to new mutations) and mutational heritability
- Past experiments typically show that most mutations are of small effects, and this is studied using the gamma distribution
- U has been estimated for multiceullular eukaryotes (U = 0.08) and viruses (0.93).
- Mean mutational coefficient of variation for fitness traits is $CV_M = 0.026$
- The Bateman-Mukai method produces estimates with bias of unknown magnitude, because we do not know the distribution of mutational effects
- rate of deleterious mutation can also be estimated from molecular data as $U = U_n\phi$, where $U_n$ is the rate of nucleotide mutation per generation for the genome and phi is the fraction of mutations whose deleterious effect is large enough for selection to determine its elimination, calculated by $\phi = 1 - \pi_g/\pi_s$ where $pi_s$ is the rate substitution rate of mutations that are assumed to be exclusively neutral, and $pi_g$ is the total divergence observed between to twin species
- Detecting mutation rates using molecular methods only discovers mutations of large effect, and those of small effect are not found because of insufficient power
- The dominance coefficient
  - Mutation accumulation experiments enable the estimates of average coefficient of dominance
  - First estimate is obtained as the ratio between the value of the trait from a cross between two lines along with the sum of the values of the trait in those lines. Mutations will be found in heterozygosis in the cross
  - The decline in mean of trait divded by the sum of declines occuring in lines makes up the ratio estimates, which is a measure of the average of dominance coefficients weighted by selection coefficients
  - Unbiased method for estimating average coefficient of dominance: regress the trait values at cross between lines of heterozygotes onto homozygotes (acting as the sum of trait values in parental lines)
  - Joint distribution of selection and dominance coefficients: hard to do because the effect of each mutation must be estimated individually
  - Large effect mutations in homozygosis are assumed to be recessive
  - Small-effect mutations are additive
  - Fisher hypothesis: a large amount of observed mutations are recessive. It is thought that they were initially additive and that dominance modifiers appeared at other loci that would cause the heterozygoye expression to become similar to wild homozygous
  - More common explanation: Wright's physiological theory of dominance. This is based on the idea that the enzymatic activity of heterozygotes is intermediate between homozygotes, but the flux of a metabolic pathway with several enzymes is non linear. But if the mutation has a high activity-reducing effect, then the heterozygote is more similar to the wild homozygote, so the mutation has recessive gene action. This explains why large-effect mutations tend to be recessive
- Beneficial mutations and summary of mutational parameters: there is a large variation in mutational parameters between estimates for different traits, species and estimation methods
- Somatic mutation rates may be 50x higher than those of the germline
- Mutations and environmental factors: can be additive or multiplicative. Can also be synergistic epistasis (greater than combining them) or antagonistic (lower than combining)
- Mutations can be conditionally deleterious (such as temperature dependent ones)
- Rate of mutation can also be contingent on the genetic background and whether a population is adapted to its environment. For instance, populations with low average fitness

- see higher rates of beneficial mutation compared to those with normal average fitness (though these are compensatory)
- Mutational parameters for quasi-neutral traits: can be studied using mutation accumulation or, starting with a population lacking genetic variability, artifical selection
- If the trait is neutral, then variability is determined exclusively by mutation (increase) or drift (reduction)
- Under the neutral model, the heterozygosity in the mutation-drift equilibrium is a function of $N_e$ and mutation rate per locus
- Alternatively, knowing heterozygosity and mutation rate we can estimate effective population size
- Additive genetic variance in each generation is the result of the decrease in drift and the increase by new mutation, so equilibrium will be reached
- Though traits are not generally neutral, the following expression can be used if $N_e$ is sufficiently small: $\hat{H} = \frac{4N_e u}{4N_e u + 1}$
- The genetic variance between mutation accumulation lines increases at a rate $2\delta F V_A$
- The response to artificial selection due to mutation can be expressed as $R_t = t\hat{V}i/\sqrt{V_P}$ where $i$ is the intensity with which selection is applied and $V_P$ is the phenotypic variance in that trait
- Genes with low effect have effect sizes that are indistinguishable from additivity, but those of large effect are usually recessive and have deleterious pleiotropic effects on fitness
- Implications of deleterious mutations in populations of large sizes: the maximum possible frequency of a deleterious allele at the mutation selection balance is $\hat{q} = u/sh$
- Small effect of alleles on heterozygosis implies a substantial reduction of equilibrium frequency
- Mutation load: decrease in mean fitness of a population compared to that with no segregating deleterious mutations
- The mutation load for a recessive allele is equal to u which implies that in deleterious mutation selection equilibrium, the load only depends on mutation rate regardless of frequency
- Joint effect of all loci with presumed deleterious alleles: $L \approx 1 - e^{-2U}$
- The mean number of mutations eliminated per individual by selection is L(z-y) where L is the load, z is the mean number of del. mutations in individuals and the y is the mean number of mutations carried by survivors

## Problems

(1) The aim is to carry out an experiment to estimate the heritability of a trait using pairs of parents and offspring. It is expected that the estimate of heritability will be around 0.6. It is intended to evaluate only one offspring per couple and a single parent or both. How many pairs of data would have to be evaluated in each case to obtain an estimate of heritability with a standard error equal to or less than 0.05?

$$SE(h^2) = \frac{2}{\sqrt{n}}$$

$$n = \left(\frac{2}{SE(h^2)}\right)^2 = \left(\frac{2}{0.05}\right)^2 = 1600$$

$$SE(h^2) = \sqrt{\frac{2}{n}}$$

$$n = \left(\frac{\sqrt{2}}{SE(h^2)}\right)^2 = \left(\frac{\sqrt{2}}{0.05}\right)^2 = 800$$

## Self Assessment

(1) The estimate of heritability obtained from the regression of offspring values on that of their parents will necessarily be biased if the latter are not a random sample of the population.

False, if the group of parents were selected, the regerssion estimates would not be affected if the bias would occur to the same extent in the numerator and the denominator of the equation $b_{OP} = \frac{cov(O,P)}{\sigma_P^2}$