

Outline of lectures 10-12

Genetics of quantitative characters

1. When we have a measurable (quantitative) character, we may not be able to discern which phenotypes correspond to which genotypes. If each genotype gives a distribution of possible phenotypes, and these overlap, the distribution of phenotypes will have peaks that correspond to the genotypic classes but these peaks can overlap, leaving us uncertain which genotype a given phenotype corresponds to.
2. As the overlap gets more extreme we can lose track of all the peaks and just see one distribution with a single peak.
3. If there are multiple loci contributing to the phenotype, the peaks also become confused with each other. Either way, we cannot do traditional Mendelian genetics, even though the underlying basis of the trait is Mendelian.
4. The traits we are talking about include not only sizes and weights: there are many others, behavioral traits (such as a measure of geotaxis), susceptibilities to disease, other measurable traits such as lymphocyte count, and even (but most significant economically) traits like racehorse speed.
5. The variation in these traits is caused by variation at the individual loci that contribute to the traits, and by variation in the environment. As in other cases, we should not try to think of whole genotypes surviving and reproducing, as they do not have offspring of exactly the same genotype, owing to Mendelian inheritance. Instead, changes in the populations are the result of changes in the gene frequencies at the individual loci. If the loci that contribute to the trait are scattered around the genome, recombination between them makes the genotype frequencies reflect the gene frequencies.
6. If the gene frequencies of the capital-letter alleles at loci A , B , C , D , and E are called p_1 , p_2 , p_3 , p_4 , and p_5 , a genotype such as $AA BB Cc dd Ee$ will have a frequency which is simply the product of the Hardy-Weinberg frequencies at each locus, which is $p_1^2 p_2^2 2p_3(1 - p_3) (1 - p_4)^2 2p_5(1 - p_5)$.
7. We are better off thinking about what forces such as natural selection (and artificial selection) do to the gene pools at individual loci, than thinking of what they do to frequencies of whole multiple-locus genotypes. There are too many of those, and since we have Mendelian genetics and not simply clonal reproduction, we can't simply ask how each genotype will reproduce itself.
8. Only if genes are closely linked does this picture need to be modified, and haplotype frequencies then have to be used.

9. Such traits have been worked on for centuries by animal and plant breeders. They apply artificial selection (usually by breeding from the best extreme of the distribution of phenotypes).
10. At the gene level, the individuals in the top end of the population are more likely to have the alleles that predispose to a large value of the character. Selecting, one changes the gene frequencies at all these loci. Random mating among the survivors, with recombination, then results in genotypes that come from these altered gene pools.
11. R. A. Fisher introduced a statistical theory of quantitative inheritance in 1918. It assumes that the trait is controlled by adding up effects of individual loci, and then adding on an environmental effect. The individual loci can have any number of alleles and any dominance pattern, and some can have much larger effects than others. But interaction between the loci is not allowed. In real life loci interact in complex ways, and environments interact with genetic effects. Fisher's theory is a rough approximation, but is surprisingly effective in practice.
12. The width of a distribution of a character is measured by its *standard deviation*. The square of this quantity is easier to use statistically. It is called the *variance*. (Fisher's 1918 paper is the first place that the word "variance" is used for this!)
13. Under his model's assumptions, Fisher showed that the variance can be broken down into three components, V_A , V_D , and V_E , the additive, dominance, and environmental variances. The environmental variance is the contribution made by the environmental variation to the variance. The additive and dominance terms are both genetic effects. The additive variance is (in effect) the effect of single copies of genes, and the dominance effect is the effect of the interaction of pairs of copies at a locus.
14. This distinction between additive and dominance variance seems very artificial. It is related to the breakdown of variance and its attribution to causes in the statistical theory of the *analysis of variance*. This is not surprising because that theory was developed by Fisher in 1922, shortly after his quantitative genetic work. In the analysis of variance (ANOVA) we have variance components due to (say) rows and due to columns, and an interaction component between them. The dominance variance is an interaction component of this sort.
15. Fisher showed how covariances (and correlations) between relatives can be written in terms of these three variance components. This in turn means that we can measure the three variance components by measuring correlations between a number of different kinds of relatives, and then solving for the values of V_A , V_D , and V_E that fit these correlations.
16. Fisher and Sewall Wright also showed how to predict the response to artificial selection from these variance components. Jay Lush of Iowa State University applied and widely publicized the resulting theory of quantitative genetics in books starting in 1937. Iowa State University became one of the chief centers developing quantitative genetics for

animal and plant breeding in the 1920's – 1950's. There was a struggle between showing breeders of the 4-H club “most beautiful calf” type and quantitative geneticists, who emphasized that the most beautiful cow did not necessarily give offspring that had the best production.

17. The fraction of the variance contributed by the V_A term is called the *heritability*, and denoted by h^2 , where

$$h^2 = \frac{V_A}{V_A + V_D + V_E}$$

(although the heritability is always written as the square of a quantity h , that is purely for historical reasons and you can ignore h itself).

18. Note that the heritability is *not* the degree of genetic variation because the V_D term is left out of the numerator. Note also that the variance components can vary from population to population, as gene frequencies and environmental regimes vary, and can change during a selection experiment. For example, if we select a line until all alleles get fixed or lost, heritability within that line will drop to zero.

19. Typically one sees response to the artificial selection. After a time one can appear to reach a “selection limit” where further response appears to have stopped. The absence of genetic variation can be tested by carrying out reverse selection or relaxed selection. It is usually helpful to also have an unselected control line, to check for environmental changes, and to have replicates of the selection lines.

- If genetic variation has been lost, we expect that reversing selection of the selected line will not get a response. Neither selection up nor down will show a response.
- An alternative possibility is that genetic variation is still present, but that the selection limit comes from natural selection which is opposing artificial selection (that the individuals you judge best have lower survival or fertility under your breeding conditions, enough to stop progress). This can be tested by relaxing selection; if there is natural selection opposing progress, it will then start pushing the population mean back towards the starting point.

20. The response to artificial selection is the heritability times the *selection differential*. The selection differential is the difference between the individuals that you select, and the original population mean. Thus if we have a population with a mean of 100 pounds, and breed from the biggest individuals such that the ones we breed from have a mean of 108 pounds, the selection differential is 8 pounds.

21. If S is the selection differential and R is the gain in one generation (the response),

$$R = h^2 S$$

This is called the *Breeder's Equation*.

22. So if the heritability is 0.4 in the above case, and R is 8 pounds, the expected gain is 3.2 pounds.

23. The reason why the additive genetic variance is the important genetic component is that individual genes get passed on, but the particular genotypes that result in dominance effects get separated by Mendelian segregation and not passed on as a package, intact. For each gamete, at a locus, one of the two copies gets into the gamete, but not the particular pair.
24. Thus the response to selection is not the full selection differential, but that part of it which reflects the contributions of the single copies of the alleles. The part of the selection differential which is achieved by having individuals with particularly good combinations of alleles is lost as Mendelian segregation separates these copies of the alleles.
25. The Fisher-Wright-Lush theory has been essential to animal and plant breeding for 75 years, but is being (*very* slowly superseded by the use of molecular markers. They can be used to map *QTLs*, “quantitative trait loci” to specific places in the genome, and we can then hope to understand them biochemically, and perhaps manipulate them genetically.
26. However for a long time it will be hard to detect loci of small effect this way, and some traits have many of them and have much of their variance contributed by them.
27. An example is human height, for which many *QTLs* been found, but these account for only about 15% of the additive genetic variation in height. The rest seem to mostly be loci of small effect, or loci in which the variation is due to rare alleles. So it may be hard to find most loci that contribute to the variation in the trait.
28. A variety of artificial selection experiments were discussed. These often involved not only “up” and “down” lines, but replicates of these lines.
 - The replicates allow us to see differences that are due to measurement error and to genetic drift in the lines.
 - The replicates will show correlated random fluctuations if those are due, not to drift or measurement error, but to fluctuations in the common environment of the experiment.
 - Occasionally, to check whether there is still genetic variation in the lines, reverse selection is tried – it will be successful in changing back in the other direction if there is genetic variability in the line.
 - We could also relax selection. Relaxation of selection leaves natural selection (which does happen in the lab) still active, and it may reverse the effects of the artificial selection.
29. Selection can be made on morphology or even on behavioral characters. Automatic self-selection machinery such as Dobzhansky’s mazes, or Ken Weber’s “inebriometer” can avoid some of the tedium of assessing individual phenotypes.

30. A notable selection experiment is the Illinois Corn Selection Experiment, carried out since 1896 at the University of Illinois. Four lines of open-pollinated (random mating) corn were selection, one each up and down in protein content and in oil content of the kernels. The down line in oil content died out, finally. But all lines continued to make progress far beyond the limits of the original population. This is most likely the result of mutations that occurred in these large populations, mutations that kept introducing variability into the populations as the original varying loci got fixed by the selection.
31. Natural selection works differently from animal and plant breeders, as it is often acting not by carrying out this kind of “truncation selection” but by selecting for an intermediate phenotype, “optimum selection”. This can be seen in studies of survival of individuals of different phenotypes in natural populations.
32. An example of this is the data collected by Hermon Bumpus in 1896, which shows the distribution of lengths (and other size measurements) of 136 House Sparrows (*Passer domesticus*) that were knocked down by a freezing storm in February, 1896. He measured 8 variables on them, and recorded which ones survived and which ones didn't. He published the whole data table. Females of intermediate length survived better (males of greater length survived better).
33. Even though the pattern of optimum selection is different from truncation selection, the theory still applies, with the response being a fraction h^2 of the selection differential.