

Chapter 12

Polygenic inheritance, quantitative genetics and heritability

So far we have considered characters determined by a single gene with two alleles, occurring in sharply contrasting states, which can have a major affect on the fitness of the organism. In some cases we are justified in modelling selection in this manner, but in many cases, probably the majority, we are not. It is possible to expand the basic theory to consider characters determined by two gene loci, but this approach is no longer useful when we consider characters that are determined by many genes. In these cases we may observe a general relationship between parent and offspring, which suggests that there is an underlying genetic basis to the trait, but we usually do not know how many genes are involved or how they interact. In addition, we may also be aware that the environment influences the trait to some extent. Consequently, in order to study these traits we examine their variability, and attempt to dissect this variation into its genetic and environmental components. This type of analysis is called quantitative genetics.

We can consider three types of quantitative traits (Hartl and Clark 1989):

1. *Meristic traits*, in which the phenotype is expressed in discrete, integral classes. Examples include litter size or number of seeds produced per individual, number of flower parts, and kernel colour in wheat.
2. *Continuous traits*, in which there is a continuum of possible phenotypes. Examples include height, weight, oil content, milk yield, human skin colour, and growth rate. In practice, similar phenotypes are often grouped together into classes for the purposes of analysis.
3. *Discrete traits*, in which an individual either does or does not express the characteristic. Multiple genetic and environmental factors combine to determine the risk or liability of expressing the trait. It is assumed that the liability has to be greater than some threshold before the trait is expressed. Examples include diabetes and schizophrenia in humans.



Fig. 12.1 Kernel colour in the F_2 generation of a cross between a white and a dark red variety of wheat. The colour difference is assumed to be due to three gene loci, and each red allele is denoted $\bullet$ and each white allele is denoted $\circ$. The 64 possible combination of alleles are grouped into the seven possible phenotypes, which occur in the proportions shown.

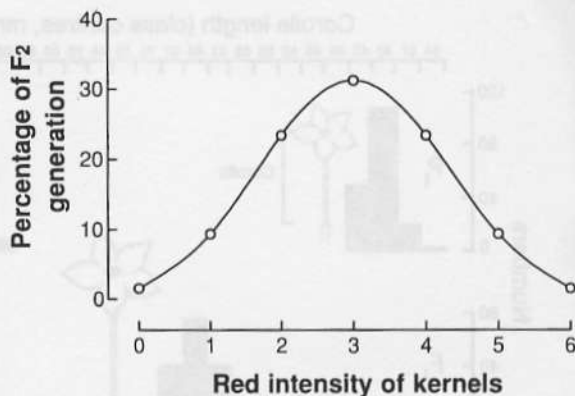
Proportions:	1/64	6/64	15/64	20/64	15/64	6/64	1/64
No. red alleles:	0	1	2	3	4	5	6
Phenotype:	White	Increasing redness				Dark red	

12.1 Polygenic inheritance

Quantitative traits are influenced by many genes, called polygenes, each one of which contributes a small amount to the variation of a character. The first genetic analysis of a quantitative trait was made by the Scandinavian geneticist Nilsson-Ehle, in 1909. He studied red versus white kernel colour in wheat, and showed that there are three gene loci governing this trait. There are red alleles (R_1 , R_2 and R_3) and white alleles (W_1 , W_2 and W_3) at each locus, and there is no dominance in their effects. The alleles act in an additive manner, so that as the number of red alleles increases the intensity of the red colour increases, or conversely as the number of white alleles increases the intensity of the red colour decreases. Nilsson-Ehle crossed a homozygous white (denoted $\circ\circ\circ\circ\circ\circ$) with a homozygous dark red strain (denoted $\bullet\bullet\bullet\bullet\bullet\bullet$) and the kernels of the F_1 were an intermediate red colour (genotype $\circ\bullet\circ\bullet\circ\bullet$). The F_1 individuals can produce $2^3 = 8$ different types of gametes, and the $F_1 \times F_1$ cross will produce $8 \times 8 = 64$ unique combinations of these alleles in the F_2 generation. As there is no dominance there are 7 possible phenotypes corresponding to 0 to 6 red alleles, which occur in a 1:6:15:20:15:6:1 ratio (Fig. 12.1).

We have considered a meristic quantitative trait in this example. It remains meristic because there is little environmental effect on kernel colour, and the alleles of the different genes act in a purely additive manner. Consequently, there are only seven discrete phenotypes. However, had the environment affected kernel colour, and if the alleles of the three different genes affected redness by slightly

Fig. 12.2 Hypothetical continuous distribution of kernel colour in wheat. The proportions of the different genotypes remain the same as in Fig. 12.1, but environmental and other genetic effects blur the distinction between different genotypes (see text).



different amounts, the boundaries between the phenotypes would become blurred so that there would be more or less a continuum in kernel colour from white to dark red. In this case, the distribution of kernel colour would follow a smooth curve (Fig. 12.2) following the general shape of the histogram in Fig. 12.1. To analyse such a continuous distribution of kernel colour we might arbitrarily group the colours into seven classes which would be related in some way to the number of red alleles per individual. Thus, we can see that there is really no distinction between the first two types of quantitative traits (i.e. meristic and continuous).

12.2 Partitioning phenotypic variation into different components

The first attempt to partition phenotypic variation into its genetic and environmental components was made by East (1916) who began his experiments on the flower length of *Nicotiana longiflora* in 1912. We will use his data in the following two subsections to show how phenotypic variation can be partitioned into its various components. Our method of analysis is kept simple for obvious reasons, and you should be aware that it is not applicable in all situations. Some of the difficulties will be briefly mentioned as we develop our analysis, but for now let us consider our use of the similarity between parent and offspring to measure the genetic basis of a trait. Behavioural traits may be genetically transmitted from parent to offspring but may also be modified or taught by the parents, and our method of analysis does not distinguish between these two modes of transmission. A more complex example is provided by the body weight of eutherian mammals. An individual's body weight at the time of weaning depends on the body weight of the parents (genetic transmission), its weight at birth and the amount of milk it receives, which are influenced by the nutritional status of the mother and the litter size or number of siblings (transmission of maternal and sibling environmental effects).

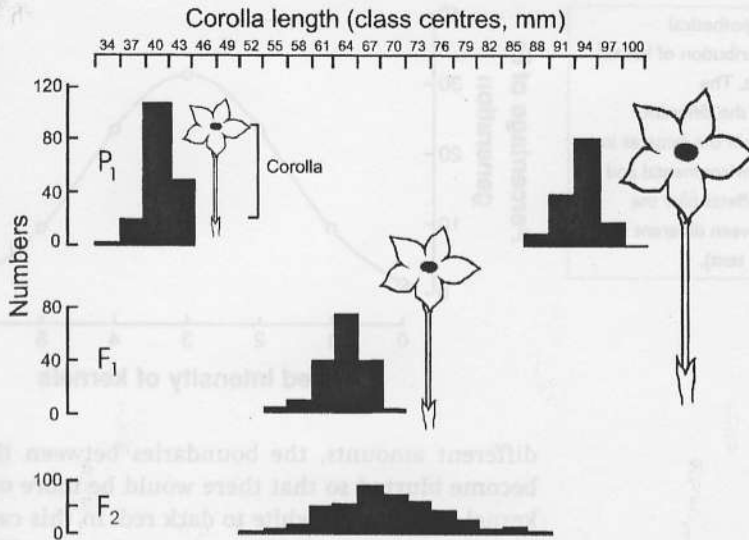


Fig. 12.3 Breeding experiments on flower length in *Nicotiana longiflora*. (Data from East 1916.)

Geneticists use a variety of methods to overcome these difficulties and more accurately partition the phenotypic variance into its different components (see Falconer and Mackay 1996), but they are more complex and are beyond the scope of this text.

12.2.1 Genetic and environmental components

The phenotypic variance can be calculated in a straightforward manner, described in any statistics text, as the average of the squared deviations about the mean phenotypic value. Phenotypic variation is divided into its genetic and environmental components by assuming that these sources of variation are additive. If this is the case, the total phenotypic variance (V_P) equals the fraction of the phenotypic variance that is a result of genetic differences between individuals (V_G) plus the fraction of the phenotypic variance resulting from differences in the environmental conditions to which individuals were exposed (V_E). Symbolically this is written:

$$V_P = V_G + V_E \quad (\text{Eqn 12.1})$$

East partitioned the variation in flower length in the following way. He crossed homozygous long-flowered plants with homozygous short-flowered plants, and the resulting F₁ plants, which were genetically identical to one another, had flowers of intermediate length (Fig. 12.3). There was no genetic variation (i.e. $V_G = 0$) in either of the parental varieties or the F₁ offspring, and so the observed variance within these groups (V_P) equals the environmental variance, V_E . The average variance of these three groups, V_E , equalled 5.2 for East's data.

East then made a cross of F₁ individuals to produce the F₂ generation. The alleles inherited from the two parental strains segregated,

and so the total phenotypic variance of the F_2 was made up of both genetic and environmental variation. The total phenotypic variance (V_P) of the F_2 offspring was 40.5. The genetic variance (V_G) can then be calculated by rearranging Eqn 12.1 as $V_G = V_P - V_E$, which gives a value of 35.3.

In summary, by analysing East's data on the phenotypic variation of flower length in *Nicotiana*, it is possible to partition the total phenotypic variation ($V_P = 40.5$) into its environmental ($V_E = 5.2$) and genetic ($V_G = 35.3$) components by assuming that these sources of variation are additive. Thus, in the F_2 generation approximately 87% of the variation was genetically based and 13% environmental based.

We can make two general points about this partitioning of phenotypic variation. First, the amount of variation (V_P) and the relative strengths of the genetic and environmental effects are not fixed entities. We may note that the value of V_G varied from zero, when the crosses were between genetically identical plants, to 35.3 for the $F_1 \times F_1$ cross, and it would be different again for crosses between other genotypes. In addition, if the plants had been grown in a more heterogeneous environment we would expect to see V_E increase for obvious reasons. Moreover, for some traits there can be genotype-environment interaction where some genotypes do better in some environments, and other genotypes do better in others. Consequently, the overall phenotypic variation and the relative importance of the genetic and environmental components vary according to the environment and the precise genetic make-up of the population.

Second, our partitioning of phenotypic variation does not give an unequivocal answer to the old genetics-versus-environment or 'nature-versus-nurture' debate. In our example of flower length it looks as though it is more important to have the 'right' genes rather than environment if we want a flower of a specific length. However, if we only had an inbred line with low genetic diversity, the reverse might be true. The debate has been highly emotional at times, and the opposing sides have often taken extreme positions, claiming either that only genetic variation is important (genetic determinism) or that the environment (nurture) is all-important. In reality it is a mixture of these two components that determines phenotypic expression, although their relative importance can vary. However, as we have seen, their relative importance is not fixed and so the debate continues without final resolution for some people. We will look at two examples of this debate in more detail, in section 12.6 of this chapter and in Chapter 19 (section 19.1).

12.2.2 Partitioning the components of genetic variation

The genetic variance (V_G) is also made up of a number of components. These components include the additive effects of all of the alleles that affect the trait, the dominance effects between alleles within gene loci, and epistatic interactions between different gene loci that

modify the additive effects. To help us understand how the additive, dominance and epistatic effects can influence the genetic variance, consider the following hypothetical series:

Genotype	aabb	Aabb	AAbb	AABb	AABB
1. Additive effect					
Phenotypic score	0	1	2	3	4
2. Dominance effect					
Phenotypic score	0	2	2	4	4
3. Dominance plus epistatic effect					
Phenotypic score	0	0	0	4	4

Imagine that this corresponds to a situation similar to that of kernel colour in wheat (section 12.1), but there are only two gene loci involved and the red alleles are represented by capital letters. When there are purely additive effects, the red colour intensifies in a stepwise fashion (0–4) as each red allele is added. Now imagine that the red allele is completely dominant to white, as shown in the second example. The intensity of the red colour would be the same whether one or both alleles of a gene coded for red, and the phenotypic scores would be modified as shown. Finally, in the third example we can imagine that the A allele only exerts its effect in the presence of allele B, and so there would be a further modification of phenotypic scores as shown.

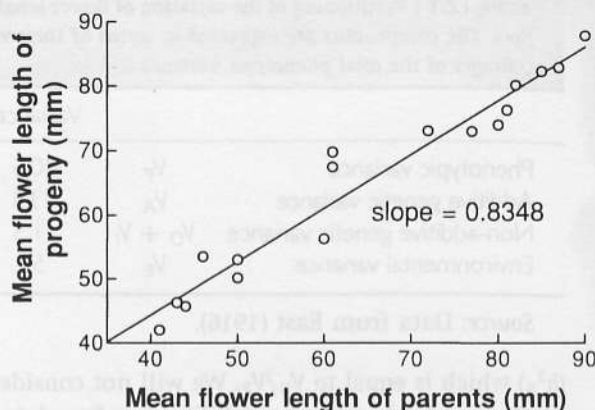
Thus, it is necessary to partition the genetic variance, V_G , into the various components as follows:

$$V_G = V_A + (V_D + V_I) \quad (\text{Eqn 12.2})$$

in which V_A is the variance due to the additive effects of alleles, V_D is the variance due to dominance effects between alleles and V_I is the variance due to epistatic interactions between the genes that affect the trait. In practice, it is difficult to separate V_D and V_I and consequently they are often grouped together as non-additive genetic variation.

The additive genetic variance (V_A) is the main cause of the resemblance between parents and their offspring, and between relatives. We can obtain a measure of this relationship by drawing a graph of the mean phenotypic score of offspring against the mean phenotypic score of their parents. Ideally, the parents should be mated at random when constructing these graphs, which can then be used to calculate V_A (see below). If we consider our example of flower length in *Nicotiana*, and use the data from crosses from the F_2 generation provided in East (1916), we obtain the following relationship between parent and offspring (Fig. 12.4).

Fig. 12.4 The relationship of flower length between parents and offspring in *Nicotiana longiflora*. (Data from East 1916.)



The slope of the regression tells us how much the offspring resemble their parents, or what is called the *heritability in the narrow sense* (h^2_N) of the trait.¹ Thus, if the offspring have the same average phenotypic score as their parents, the slope of the regression (h^2_N) will be 1.0, and if there is no relationship in the phenotypic scores of parents and their offspring, then $h^2_N = 0$. Obviously, the higher the heritability (or slope of the regression) the larger the additive genetic component. The relationship between heritability (h^2_N), additive genetic variance (V_A) and phenotypic variance (V_P) is given by:

$$h^2_N = \frac{V_A}{V_P} \quad (\text{Eqn 12.3})$$

From East's data (Fig. 12.4) we see that $h^2_N = 0.8348$ for flower length in *Nicotiana*. In section 12.2.1, we noted that $V_P = 40.5$, and so we can estimate V_A as $0.8348 \times 40.5 = 33.8$ by rearranging Eqn 12.3. We have previously estimated the genetic variance (V_G) as 35.3, and so from Eqn 12.2 we can estimate the non-additive genetic variance ($V_D + V_I$) as $35.3 - 33.8 = 1.5$.

This completes our partitioning of the phenotypic variation into its various genetic and environmental components, and the results are summarized in Table 12.1. The genetic variance (V_G) is the sum of the additive and non-additive genetic variances, and equals 35.3, or 87% of the total phenotypic variance.

12.3 Heritability

We have just seen that heritability in the narrow sense (h^2_N) is the proportion of the total phenotypic variation that is a result of additive genetic variation (Eqn 12.3). You should also be aware that there is another measure of heritability, called *heritability in the broad sense*

¹ The degree of genetic determination, or heritability, of a trait is symbolized as h^2 because it was first calculated as the square of the partial correlation coefficient (i.e. path coefficient) between the parental genotypes and the offspring's phenotype (see Feldman 1992).

Table 12.1 Partitioning of the variation of flower length in *Nicotiana longiflora*. The components are expressed in terms of their variance and as percentages of the total phenotypic variance

		Variance	Percentage
Phenotypic variance	V_P	40.5	100
Additive genetic variance	V_A	33.8	83
Non-additive genetic variance	$V_D + V_I$	1.5	4
Environmental variance	V_E	5.2	13

Source: Data from East (1916).

(h^2_B) which is equal to V_G/V_P . We will not consider this measure any further, and wherever heritability is referred to in this chapter it means heritability in the narrow sense.

The term heritability has unfortunate connotations, and is frequently misunderstood, particularly by non-biologists. Many people believe it is a fixed property for a particular trait, and think that a character is genetically determined to a certain extent and is modified by the environment by some other, usually small, amount. This is not the case. Heritability is simply a ratio of two variances, and is only applicable to the population and environment in which it was measured. We can understand this if we expand Eqn 12.3 to:

$$h^2_N = \frac{V_A}{V_G + V_E} \quad (\text{Exp. 12.1})$$

The value of h^2_N is changed if we change the genetic constitution of the population because the variance of at least one of the genetic components will be altered. For example, if we had estimated the heritability of flower length for either of the two parental populations of *Nicotiana* we would have obtained values of 0 (zero), instead of the value of 0.8348 estimated in section 12.2.2. This is because there is no genetic variation (V_G and $V_A = 0$) in these two homozygous populations, and all of the variation is a result of environmental variation (V_E). Similarly, changes to the environment can also change the value of h^2_N . For example, height might have a high heritability for a population of plants grown under very uniform conditions, but if we grew the same genetic stock in an area where the soil and water conditions were extremely variable, the heritability would be lowered because the environmental variance (V_E) would increase.

Bearing this in mind when we compare the heritabilities of different characteristics, we find that the heritability of trivial, apparently unimportant characteristics is frequently high, whereas the heritability is usually low for characteristics that are closely related to fitness (Table 12.2). This is because selection on trivial characters will probably be low or non-existent, and so natural selection tolerates large genetic variability in these characteristics. However, there will be strong selection pressures on traits that play a vital role in the fitness of an organism, and so generally there will be much less genetic

Table 12.2 Approximate values of the heritability of various characters in certain domestic animal and plant species. Traits closely related to fitness (e.g. calving interval, eggs per hen, litter size of swine, yield and ear number of corn) tend to have low heritabilities

Species and trait	h^2_N
<i>Cattle</i>	
Wither height	0.60
Milk protein percentage	0.55
Feed efficiency	0.35
Milk yield	0.30
Calving interval	0.25
<i>Poultry</i>	
Egg weight	0.55
Body weight	0.50
Albumen content	0.40
Age of sexual maturity	0.35
Eggs per hen	0.10
<i>Swine</i>	
Back-fat thickness	0.60
Body length	0.53
Feed efficiency	0.35
Daily gain in weight	0.30
Litter size	0.15
<i>Corn (Zea mays)</i>	
Husk extension	0.67
Plant height	0.53
Ear height	0.45
Ear number	0.20
Yield	0.13

Source: Data from Hartl and Clark (1989).

variation because the inferior genotypes will be eliminated from the population.

Plant and animal breeders are interested in the heritabilities of different characteristics because the higher the heritability, the greater the response to selection. This leads us to our next topic where we consider the effect of selection on quantitative characters.

12.4 Response to selection

How do quantitative characters respond to selection? In many cases they will change, and we can illustrate this over two generations of selection using an abstract example (Fig. 12.5). The phenotypic score is arbitrary, and could correspond to such traits as the amount of oil in a seed, plant height, the degree of resistance to a particular insecticide, or body weight. We apply systematic selection to increase the size of the character in question. In the original population

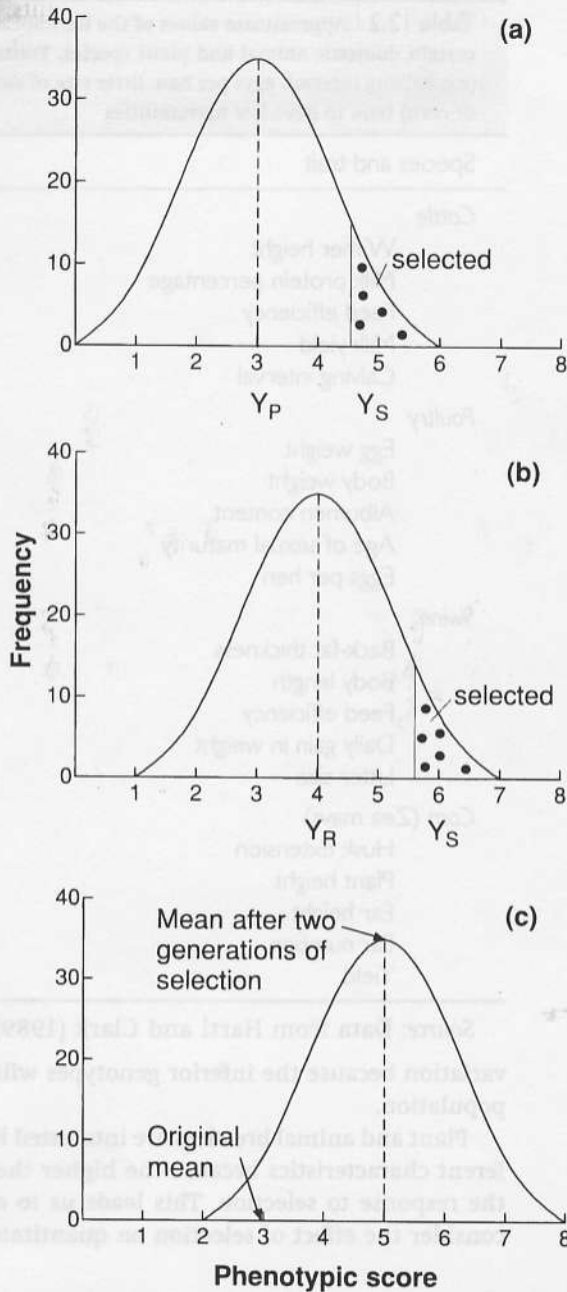


Fig. 12.5 Two generations of selection for increased size of a trait with a heritability of 0.5. The individuals selected to be the parents of the next generation are stippled (see text).

(Fig. 12.5a) we can see that the overall phenotypic mean of the parental population ($\bar{Y}_P$) is 3 units, and the group of individuals selected as parents of the next generation have an overall mean ($\bar{Y}_S$) of 5 units.

The intensity of selection, or selection pressure, being applied is called the *selection differential* (S), and is measured as the difference between the mean of the selected parents ($\bar{Y}_S$) and the mean of all the individuals in the parental population ($\bar{Y}_P$). Thus:

$$S = \bar{Y}_S - \bar{Y}_P$$

(Eqn 12.4)