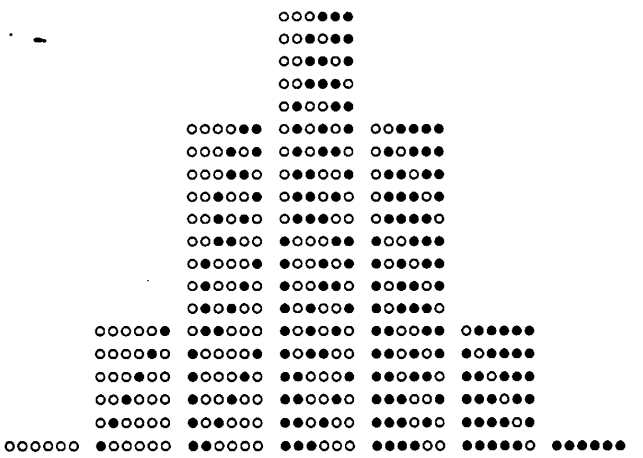
Chapter 12

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.



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

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 • and each white allele is denoted o. The 64 possible combination of alleles are grouped into the seven possible phenotypes, which occur in the proportions shown.

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 oooooo) with a homozygous dark red strain (denoted ••••••) and the kernels of the F_1 were an intermediate red colour (genotype o•o•o•). 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

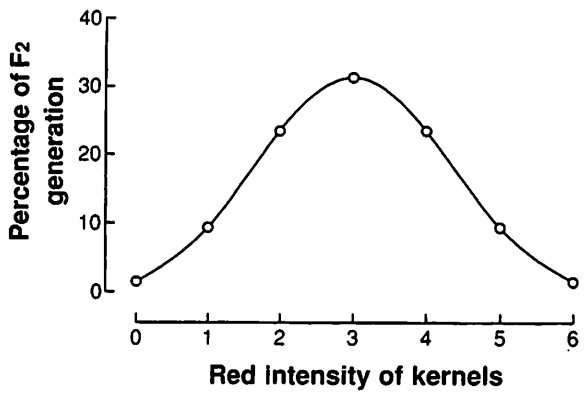


Fig. 12.2 Hypothetical continuous distribution of kernel colour in wheat. The proportions of the different genotypes remain the same as in Fig. 12.1, but environmental and other genetic effects blur the distinction between different genotypes (see text).

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).

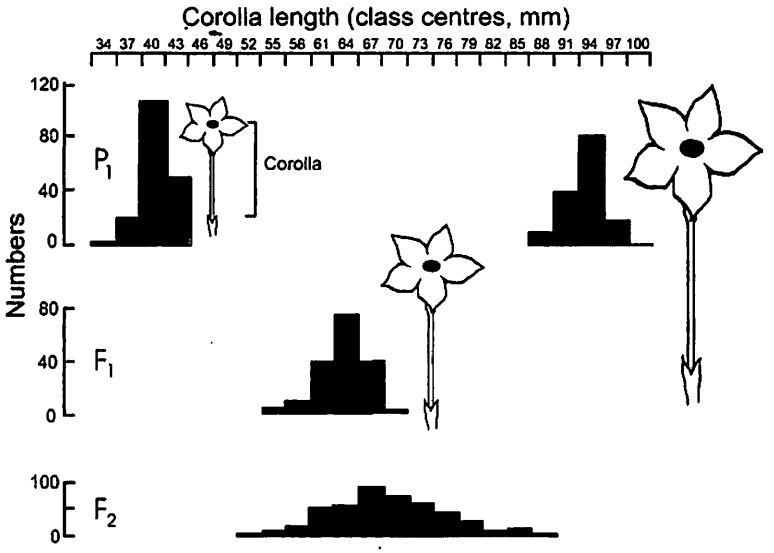


Fig. 12.3 Breeding experiments on flower length in *Nicotiana longiflora*. (Data from East 1916.)

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:

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

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 F_1 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 F_1 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 F_1 individuals to produce the F_2 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:

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

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).

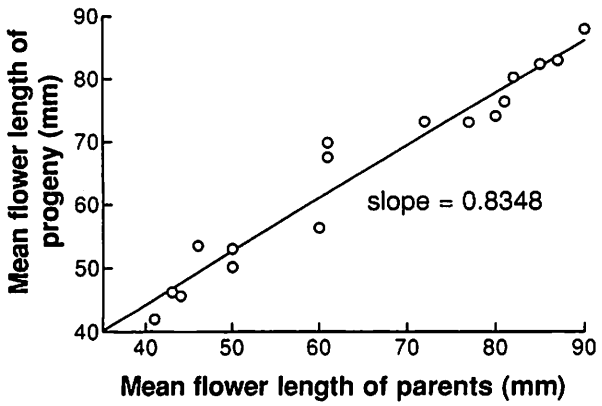


Fig. 12.4 The relationship of flower length between parents and offspring in *Nicotiana longiflora*. (Data from East 1916.)

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:

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

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:

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

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
Cattle	
Wither height	0.60
Milk protein percentage	0.55
Feed efficiency	0.35
Milk yield	0.30
Calving interval	0.25
Poultry	
Egg weight	0.55
Body weight	0.50
Albumen content	0.40
Age of sexual maturity	0.35
Eggs per hen	0.10
Swine	
Back-fat thickness	0.60
Body length	0.53
Feed efficiency	0.35
Daily gain in weight	0.30
Litter size	0.15
Corn (Zea mays)	
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

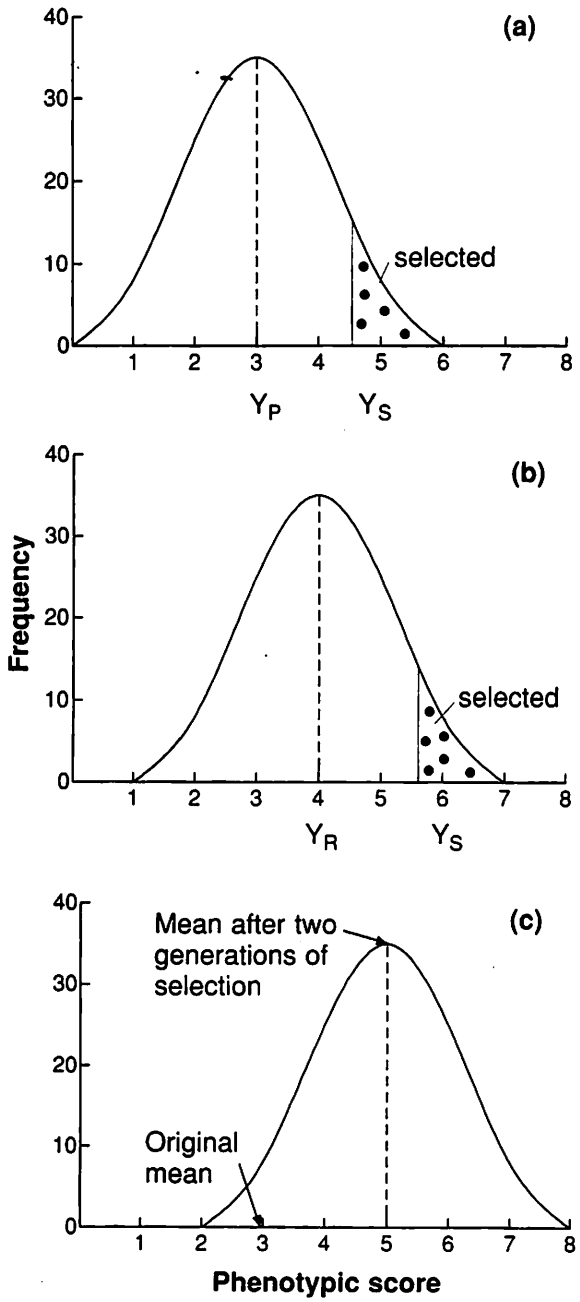


Fig. 12.5 Two generations of selection for increased size of a trait with a heritability of 0.5. The individuals selected to be the parents of the next generation are stippled (see text).

(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:

$$S = \bar{Y}_S - \bar{Y}_P \qquad \text{(Eqn 12.4)}$$

In our example we can see that S = 5 − 3, or 2 phenotypic units.

The response (R) to this selection differential is shown in Fig. 12.5b, and may be measured as the difference between the mean of the progeny ($\bar{Y}_R$) and the mean of the parental population ($\bar{Y}_P$). Thus:

$$R = \bar{Y}_R - \bar{Y}_P \qquad \text{(Eqn 12.5)}$$

In our example, $R = 4 - 3$, or 1 phenotypic unit.

We can also predict the response (R) of the population from the equation:

$$R = h^2_N S \qquad \text{(Eqn 12.6)}$$

On reflection, this last equation should be intuitively obvious. The heritability (h^2_N) is a measure of the similarity of the progeny to their parents, and if we multiply this by the selection differential (S) we can predict the overall change in the phenotype (or response, R). In our example, $R = 0.5 \times 2$, or 1 phenotypic unit, which is the same as that predicted by Eqn 12.5. Thus, the overall phenotypic score will change, or in this case increase because R is positive, by 1 phenotypic unit. The average phenotypic score of the progeny should be one unit larger than the mean score of the parents, i.e. 3 + 1 = 4, which is what we observe.

When we repeat the operation, we see that after a second generation of selection at the same intensity ($S = 2$), the average phenotypic score of the population has increased by one more unit and now is 2 units larger than the original population (Fig. 12.5c). Thus, the population responds to selection generation after generation, and the response is directly related to the heritability (h^2_N) and the selection differential (S) being applied. If the heritability had been lower, the response to selection would have been lower. For example if h^2_N was 0.25 and $S = 2$ units, the mean phenotypic score would have increased by 0.25×2 or 0.5 units each generation, rather than the 1 unit observed in Fig. 12.5. Note, that unless the heritability = 1.0, the mean phenotypic score of the progeny will lie somewhere between the phenotypic score of those selected as parents ($\bar{Y}_S$) and the overall phenotypic score of the parental population ($\bar{Y}_P$). This slipping back toward the overall mean is known as regression (toward the mean).

12.5 Empirical examples of selection of quantitative characters

Perhaps the best example of controlled, long-term selection is an experiment on the oil content of corn seed. The experiment was started in 1896 (even before the rediscovery of Mendel's laws) at the Illinois Experimental Station using a base population of 163 corn ears with oil contents ranging from 4% to 6%. Two experimental lines were started, one selecting for high oil content and the other for low oil

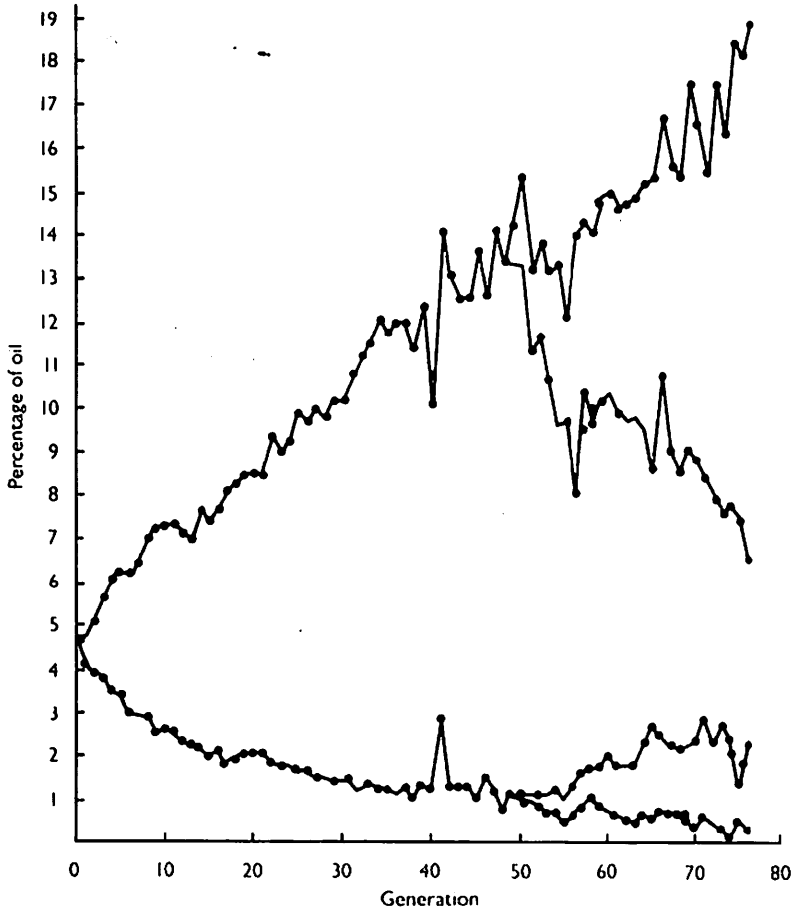


Fig. 12.6 Selection for high and low oil content in the seed of corn (*Zea mays*). The high line was formed from the 24 ears highest in oil content and the low line from the 12 ears with the lowest oil content in an initial population of 163 corn ears. Two additional lines of reverse selection were started in generation 46. (Reprinted from Dudley, J. W., in *Proceedings of the International Conference on Quantitative Genetics*) (eds. E. Pollock, O. Kempthorne and T. B. Bailey), pp. 459–73, Copyright 1977, with permission of Iowa State University Press.)

content and the results of the first 78 generations of selection are shown (Fig. 12.6).

It did not take long for the oil content of the two lines to be well outside of the original range of oil content. The low line has changed less than the high line because you cannot have an oil content of below zero. A careful examination of the trends shows that the rate of change was fastest during the first 10 to 20 generations of selection, but the lines are still diverging and so the additive genetic variance has still not been exhausted. Presumably, at some stage all the individuals within a line will come to have the same genotype for oil content, and the heritability (h^2_N) will be zero because there is no additive genetic variance (V_A). At this point there will be no further response to artificial selection unless there are new mutations affecting oil content. Heritability has declined to about one-third of its initial value in both lines (Table 12.3), but will take some time to decline to zero because there are at least 20 genes that affect oil content.

Selection is not always directional, it is often against the phenotypic extremes and intermediate phenotypes have higher fitness. Such selection is called stabilizing selection, because no change may

Table 12.3 The heritability of oil content in corn seed after different numbers of generations of selection in the Illinois corn experiment illustrated in Fig. 12.6

	Heritability of oil content	
Generations	High line	Low line
1 – 9	0.32	0.50
10 – 25	0.34	0.23
26 – 52	0.11	0.10
53 – 76	0.12	0.15

Source: Reprinted from Dudley, J. W. in *Proceedings of the International Conference on Quantitative Genetics* (eds. E. Pollock, O. Kempthorne and T. B. Bailey), pp. 459–73, Copyright 1977, with permission of Iowa State University Press.

be seen in the phenotypic score generation after generation, even though there may be strong selection pressures on the character in question. A good example of this type of selection is seen in human birth weights (Fig. 12.7), where there is strong selection against large and small babies and the average birth weight is similar to the optimum predicted from infant mortality. A similar pattern of stabilizing selection is seen in clutch size in birds (see Chapter 16) where intermediate clutch sizes give rise to largest number of survivors.

12.6 Intelligence, race and societal class

In recent decades there has been a controversy about the reasons for differences in IQ (intelligence quotient) scores between different segments of society, particularly in North America. For example, Jensen (1969) and Herrnstein and Murray (1994) have presented overwhelming evidence that the average IQ scores of people of different colour, and of different socioeconomic classes, are very different. The issue they raise can be stated quite simply: does the low IQ in some racial groups, or in some lower socioeconomic classes, have a genetic or an environmental basis? Thus, we are examining an example of the classic nature (i.e. genes) versus nurture (i.e. environment) debate. In the following discussion we will concentrate on racial differences, because the arguments apply equally to socioeconomic class differences and, in any case, there is frequently a strong association between race and socioeconomic class.

How do we define intelligence? It is highly questionable whether IQ scores give an unbiased and total assessment of intelligence because some things, such as musical skills, are not correlated with IQ. However, IQ tests do evaluate a large range of skills, and the scores correlate reasonably well with the scholastic success of Caucasians, and

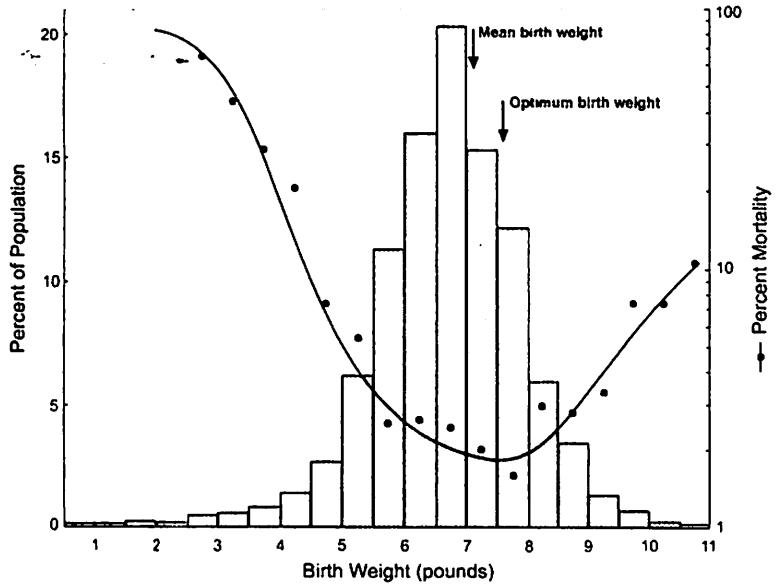


Fig. 12.7 Stabilizing selection for human birth weight (histogram). Early mortality, shown by the points around a fitted curve, is lowest near the mean birth weight. (From *The Genetics of Human Populations*, by L. L. Cavalli-Sforza and W. F. Bodmer © 1971 by W. H. Freeman and Company. Used with permission.)

so they may serve an educational purpose. The form of the present IQ tests was developed soon after 1900 by Alfred Binet, a French psychologist, to identify schoolchildren who would be likely to have difficulty at school. The Binet test was revised in 1916 at Stanford University, and subsequently revised again in 1937 and 1960. It is the standard IQ test in North America.

Interestingly, the 1916 Stanford–Binet test indicated that boys had higher IQs than girls. This was interpreted as a bias favouring males, because of their different experiences, rather than any innate tendency for males to be more intelligent than females. The test was revised to exclude questions where one sex or the other performed better, and so present tests give almost identical IQ distributions for males and females. It is interesting that there has been no successful attempt to eliminate questions where there are different scores for other different racial, cultural or socioeconomic groups.

The test involves answering a large number of questions. The answers are scored and the scores are summed for each individual. These scores are transformed into an IQ test score for each individual by a process of standardization, in which the mean for a large population has a value of 100 and the standard deviation for the population is 15. Thus, the population has a normal distribution of IQs, such that 67% of the population has IQ scores between 85 and 115, 95% of the population has IQ scores between 70 and 130, and 99% of the population has IQ scores between 55 and 145. Naturally, different tests are given to children of different ages. The results of different tests over short time intervals are similar, but the correlation between tests declines as the time interval increases. Some studies have shown that individual IQs vary little between the ages of 2 and 18, but other studies show that

they vary over this time period. For example, one study of Swedish schoolchildren showed that the IQs of students tested at 12 years of age increased by an average of 11 points at the end of high school, if they had been subjected to a demanding curriculum. Thus, IQ scores are not perfectly stable with age, and can be increased with appropriate schooling.

There have been many studies to determine the heritability of IQ scores to assess the genetic and non-genetic aspects of this measure. However, there are problems if one uses the parent–offspring regression technique (see Fig. 12.4) because relatives reared together cannot provide data that discriminate between genetic transmission of the trait and environmental transmission from parent to child. These estimates of heritability are likely to be too high because 'good' genes and 'good' environments are likely to be highly correlated with one another. Better estimates can be made by using the technique of estimating the correlations between monozygotic twins that have been reared apart.[2] There are four such studies providing estimates of 0.86 (Cyril Burt), 0.77 (J. Shields), 0.73 (Juel-Nielson), and 0.69 (Newman, Freeman and Holzinger). These give a combined estimate of 0.81, which is where the widely quoted estimate of 80% for the heritability of IQ has been derived. However, we now know that Burt's data are fraudulent (he made up most of his data) and Shield's data are less than ideal because over two-thirds of his sample were raised by close relatives. The remaining two estimates are based on small sample sizes, and so their combined estimate of 0.71 is rather uncertain. Finally, studies based on the results of adopted children have provided estimates of between 0.45 and 0.65. Thus, the claim that IQ is 80% heritable is certainly an inflated estimate, and more reasonable estimates would be somewhere in the range of 45% to 70%, or even lower (Feldman 1992). However, as we will see, the precise range of heritability estimates for IQ does not impair our ability to evaluate whether the differences between groups have a genetic or an environmental basis.

Numerous studies have shown that the average IQ of blacks in the USA is 10 to 20 points (average 15 points) below that of US whites (Fig. 12.8). If we ignore the question of how the authors of these studies defined 'white' and 'black' people, this represents a considerable difference between the two groups, equivalent to one standard deviation of the distribution of white IQs. Similarly, the average IQ of people in lower socioeconomic classes is below that of people in higher economic classes.

How do we explain these differences? A group comprising mainly psychologists, called hereditarians, argues that the difference between groups is largely genetic in origin. In their view, one's IQ is mainly innate, and no amount of environmental change will

[2] Monozygotic twins may be defined as: two individuals that develop by the division and separation of a single fertilized egg into two genetically identical parts (i.e. identical twins).

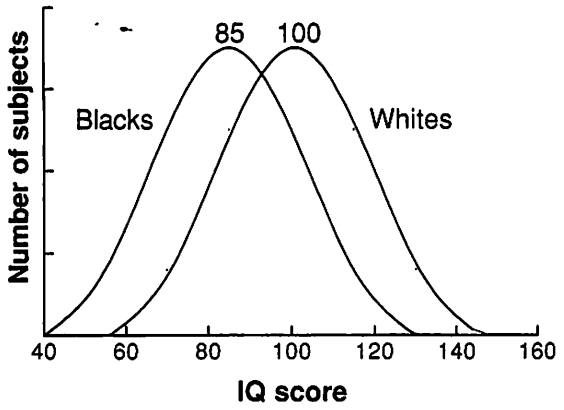


Fig. 12.8 The difference in IQ between idealized samples from blacks and whites in the United States. The mean IQ is shown above the curves for each of the two groups.

eradicate the difference between groups. In case you think that this is merely a racist argument for the genetic superiority of Caucasians, you should be aware that at least one hereditarian, William Rushton of the University of Western Ontario, shows that the average IQ of people of Asian origin is higher than that of Caucasians. On the other hand, many biologists have argued that the differences are mainly due to environmental factors, together with cultural biases in the IQ test questions that favour whites.

The hereditarians' main argument is that the difference in IQ between the two racial groups is mostly genetic in origin, because IQ has a high heritability in both blacks and whites. In other words, hereditarians believe that IQ is mainly genetically determined (as much as 80%), and so this must be the reason for the difference between the two groups. This is a fallacious argument as the following 'thought experiments' make clear.

Imagine we have a population of dogs which vary greatly in size, and may be black, grey or white (the black and white alleles are codominant). Note that this imaginary genetic system for body colour is simply to make it easy to create pure-breeding groups of black or white dogs. Dog size is a polygenic trait with a heritability of 0.8, and there is no association between dog size and colour. In the first experiment we take a group of young white puppies and feed them on a rich diet. At adulthood, these dogs have a body length that is normally distributed with a mean of 100 cm and a standard deviation of 15 cm. The same is true with the offspring of these dogs. We also take a group of young black puppies and feed them on a reduced diet. At adulthood, these dogs have a body length that is also normally distributed, but with a mean of 85 cm and a standard deviation of 15 cm. The same is true for their offspring. The slope of the parent–offspring regression for body length was 0.8 ($= h^2_N$) in both groups, but the regression line for black dogs was below that of the white dogs. Obviously, we would attribute the difference in size between the two groups of dogs to the difference in their diets, i.e. environmental factors, even though dog size has a high heritability. Thus, even though the variation in dog size *within* each group was

mainly due to additive genetic effects (i.e. V_A is 80% of V_P), this was not the reason for the difference *between* the two groups. We can also note that dog size had nothing to do with their colour, even though white dogs were significantly larger than black dogs.

In the second imaginary experiment, we select a group of white dogs for increased size, and simultaneously select a group of black dogs for smaller size. After a few generations of selection we observe that the two groups have diverged in average size. We then take the two groups and grow them in the same area, thereby eliminating potential environmental differences between them. We are not surprised that the white dogs are larger on average than the black dogs, and would attribute this difference to genetic differences in the two lines. As a result of the selective process, we might find that the heritability of dog size was reduced to approximately 0.3 in each of the two lines (see Table 12.3). Thus, even though the variation in dog size within each group was mainly a result of a combination of non-additive genetic effects and environmental effects, the difference between the two groups had a genetic basis.

These imaginary experiments demonstrate that there is no foundation to the hereditarian argument: that if IQ has a high heritability, the differences in IQ scores between different population groups must have a genetic basis. However, even though we have shown that the main hereditarian argument is based on a fallacy, this does not settle the question as to whether the difference in IQ between groups has an environmental or a genetic basis. This requires more direct evidence.

Three types of evidence suggest that most of the difference between groups is a result of environmental factors.

1. Studies that compare the IQ of different races usually have poorly matched samples. The samples not only differ in race, but also differ in socioeconomic status, the type of schooling available to the children, and so on. Thus, differences between groups could be a result of associated environmental differences rather than racial differences. Nichols and Anderson (1973) reduced some of the environmental differences in their study by comparing different ethnic groups of children whose parents had the same occupations, education and income (socioeconomic status). They tested two samples of seven-year-old children of blacks and whites in the United States: one group of high socioeconomic status from Boston, and one group of relatively low socioeconomic status from Baltimore and Philadelphia. The Boston whites and blacks had mean IQs of 104.2 and 100.0, respectively, and the Baltimore–Philadelphia whites and blacks had mean IQs of 95.3 and 91.2, respectively. Thus, the 10 to 20 point difference in IQ between blacks and whites, found in most studies (see above), was effectively reduced to about 4 points simply by stratifying the sample design so that potential racial differences were not confounded by differences arising as a result of socioeconomic factors.

2. If the hereditarian view is correct, children of black and white parents should have IQs intermediate between the two groups, but this is not the case. There was no difference in the average IQ of illegitimate children of black and white American servicemen and German women after the Second World War. Similarly, Tizard (1974) found that mixed-race and white English children who had spent part of their early life in institutions had similar IQ scores, although those of the mixed-race children were marginally higher. Finally, the IQ of children of mixed marriages is reported to correlate more with the IQ of the mother than the father, suggesting that cultural traits transmitted to the child during infancy and early childhood are important, because it is the mother who normally has the major influence on the child at this stage.
3. Evidence for environmental effects transmitted by the adoptive parents is also provided by children adopted soon after birth. The effects are quite general and apply equally to the question of racial differences in IQ as well as differences in IQ related to socioeconomic class. A study by Skodak and Skeels (see Bodmer and Cavelli-Sforza 1976) tested the IQs of 100 children, born to white, unmarried mothers from lower classes, that were adopted into middle-class families. The children's IQs showed a higher correlation with those of the biological mothers than with their adoptive parents, showing a strong genetic effect on IQ. However, the average IQ of the biological mothers was 85 whereas the average for the children was 107. The IQ of the biological fathers was unknown, but it is unlikely that all of the 22 point difference was from that source. Probably the increase in IQ was related mainly to the favourable academic environment offered by the foster homes. Finally, a study of black children adopted by white families in Minnesota showed that they had IQs similar to those of white controls, which suggests that there are no genetic differences in IQ between the races.

From this discussion, we can conclude that the hereditarian argument, which states that differences in IQ between different population groups is mainly a result of genetic differences among groups, is based on a misunderstanding of quantitative genetic theory. The available evidence suggests that most, if not all, of the differences among groups are a result of environmental influences.

What if we had discovered that much of the difference in IQs between groups had a genetic basis? Would the genetic fatalism of the hereditarians, and the belief that group differences in IQ cannot be changed, be justified? The simple answer is no. Consider the case of phenylketonuria, a genetic disease which, among other effects, causes severe mental retardation. At one time there was no treatment for the condition, and in the 1960s it accounted for almost 1% of all severely retarded patients in institutions. Phenylketonuria is a metabolic defect resulting from the absence of an enzyme called phenylalanine hydroxylase, which converts the amino acid phenylalanine to

tyrosine. The phenotypic effects of the condition, only expressed in homozygous recessive individuals, are the result of an excess of phenylalanine and its derivatives. The genetic condition can be treated by restricting the dietary intake of phenylalanine. Individuals that are identified soon after birth with the condition, before the classical phenotypic effects are expressed, and who adopt a rigorously controlled diet can live a normal life. Treatment will not reverse the effects of the disease once they have been expressed. Note, however, that phenylketonuria has a close to 100% heritability, and yet an environmental change (of diet) can effectively abolish its effect on IQ.

Finally, my vigorous attack on the hereditarian position of ascribing the differences in average IQ between racial groups to genetic differences may have led some to conclude that there is no genetic basis to IQ and that only the environment is important. However, although appropriate mental stimulation is necessary to develop our IQ to its full potential, some of us will have a higher IQ than others. We cannot all be as clever as Albert Einstein or Madame Curie, our genetic constitution is also important.

12.7 Summary

When a trait is determined by many gene loci there may be a continuum or a very large number of possible phenotypes. The phenotypic variation of polygenic traits can be partitioned into environmental and genetic components, and the latter further partitioned into additive, dominance and epistatic genetic effects, by suitable breeding experiments.

The response to selection is positively correlated with the proportion of the overall phenotypic variation (V_P) that is determined by the additive genetic effects (V_A). The measure V_A/V_P is called the heritability in the narrow sense of a trait, and when a trait is subjected to intense selection its heritability decreases as inferior alleles are eliminated. This probably explains why traits that are closely related to fitness tend to have low heritabilities, whereas trivial characters tend to have high heritabilities.

The relationship between IQ and race is examined in some detail as an example of the genes versus environment debate. The hereditarian argument that the differences in IQ between groups must have a genetic basis (because the heritability of the trait is high) is shown to be based on a misunderstanding of quantitative genetic theory. The available evidence suggests that most, if not all, of the differences between groups are the result of environmental influences. However, the genetic constitution of individuals has a strong effect on IQ.

12.8 Problems

1. Two highly inbred lines of mice are crossed and give a variance of 2.5 units in tail length in the F_1 generation. The variance in the F_2 generation is 12.5 units, and the parent-offspring regression for tail length is 0.75.

(a) What is the heritability (h^2_N) for tail length?

(b) Partition the variance (V_P) of the F_2 generation into the following components: V_E, V_G, V_A and (V_D+V_I).

(c) If the overall average tail length of hybrids is 80 cm, what will be the average tail length of the offspring of parents with tail lengths of 88 and 92 cm? What would your answer be if the average tail length of the parents had been 70 cm?