Quantitative genetics

Nature and nurture

The subject of this chapter is the inheritance of traits that are influenced by genes at many loci—that is, of polygenic inheritance. Such traits are also influenced by the environment. Of course, a phenotype that can be caused by a single mutation may also be caused by a specific environment: for example, a fruitfly may lack a particular crossvein in the wing because it is homozygous for the mutant crossveinless, or because it was exposed to a heat shock as a pupa. The 'nature-nurture' problem is discussed here, however, because the analysis of causation becomes difficult for polygenic traits.

A difference between two individuals may be genetic or environmental; that is, it may be caused by differences between the genes present in the fertilized eggs from which they developed (that is, by nature), or by differences between the environments in which they were raised (that is, by nurture). To a geneticist, any difference that is not genetic in the above sense is environmental. The reason for treating this distinction as fundamental is that, unless Lamarckism is true, only genetic differences will influence the nature of the progeny. Of course, children may resemble their parents because they share a common environment, and not only because they share genes. You will notice that, in many of the models discussed in this chapter, it is explicitly assumed that genetic and environmental factors act independently; that is, relatives do not share a common environment. Models which allow for shared environments, or for the fact that traits acquired by a parent may be transmitted culturally to the children, are necessarily more complex.

However, the definition does lead us to lump together as environmental several distinct kinds of difference:
1. Differences caused by external environmental conditions—for example temperature or nutrition.

2. Differences due to developmental noise. Figure 6.1 shows the number of abdominal bristles in an isogenic line of *Drosophila*. The members of an isogenic line are genetically very similar to one another, yet, even if raised in as uniform an environment as possible, they differ in phenotype, often to a marked degree. It is possible that these differences are caused by minor and uncontrollable differences in the external environment, but it is more likely that they arise from chance internal events during development.

3. Cytoplasmic effects. There may be a difference in non-chromosomal DNA—for example, mitochondrial or chloroplast DNA. Such differences can have long-term evolutionary consequences, although the pattern of inheritance is different. They are ignored in this chapter, and discussed on pp. 153–6. Other cytoplasmic effects occur, but are much less stable; they too are ignored in this chapter.

![Figure 6.1: Variability in an isogenic line of *Drosophila melanogaster*. Numbers of males with different numbers of bristles in a population made isogenic for all chromosomes by the technique illustrated in Fig. 4.3 (data from Dr K. Fowler).](image)

Usually, both genetic and environmental causes of variation are present simultaneously. If so, the first question is ask is whether they act *additively*. Thus consider the two sets of data in Table 6.1. Scottish flies are larger than those from Israel, and flies raised at 15 °C are larger than those raised at 25 °C. These two effects act additively, in the sense that the effect of temperature is almost the same in the two populations (20.4 units in one population, and 20.6 in the other), and the effect of genotype is the same at the two temperatures (9.9 units at 15 °C and 10.1 units at 25 °C). Additivity implies that the joint effect is the sum of the separate effects.
Contrast this with the data on growth rate in mice. Strain A grows faster with good nutrition, but slower with bad nutrition. The effects of genes and environment are no longer additive.

Now suppose that we have a single, genetically variable population, living in a range of environments. The phenotypic variability of the population for some trait can be measured by its variance,

$$V = \frac{1}{n} \sum (x_i - \bar{x})^2,$$  \hspace{1cm} (6.1)

where $x_i$ measures the phenotype of the $i$th individual, $\bar{x}$ is the mean value, and $n$ the number of individuals. If genetic and environmental factors act additively, as in the example of wing length in Drosophila, and if there is no association between the genotype of an individual and its environment, then

$$V = V_G + V_E,$$ \hspace{1cm} (6.2)

where $V_G$ and $V_E$ are the genetic and environmental variances, respectively. A third term, $V_G E$, is needed if there is gene–environment interaction, as for growth rate in mice. When there is such interaction, it is useful to think of the norm of reaction of a genotype: this is the set of phenotypes produced by the genotype in different environments.

The important points made in this section are

1. Differences can be genetic or environmental: only genetic differences will affect the nature of the offspring.
2. Causes may act additively or non-additively: if causes act additively, the effect of cause A is the same, whether or not cause B has also acted.

The additive genetic model

In this section, I work out some of the consequences of a simple model, which has two main assumptions:

(1) all differences are genetic; and
(2) genes act additively.

The assumption about additive gene effects has two parts, concerning within-locus and between-locus effects. Within a locus, it assumes that the phenotype of the heterozygote, $Aa$, is exactly intermediate between the homozygotes, $aa$ and $AA$. Thus the effect of introducing the first $A$ allele ($aa \rightarrow Aa$) is the same as the effect of introducing the second allele ($Aa \rightarrow AA$). Between loci, it assumes that the effect of a gene substitution at one locus is independent of what alleles are present at a second locus. Thus, in a haploid the difference between $ab$ and $aB$ is the same as the difference between $Ab$ and $AB$. Within a locus, the alternative to additivity is dominance: between loci, it is epistasis. In terms of variance, therefore, we can write

$$V_G = V_A + V_D + V_I,$$

where $V_A$ = additive genetic variance; $V_D$ = dominance variance; and $V_I$ = variance due to epistasis (= interaction between loci).

The reason for singling out the additive genetic variance for special attention is that, as we shall see, it is the component of the total variance that causes the response of a population to selection, natural or artificial.

Of course, our model is not true of real populations. The environment does affect the phenotype, and genes do not always act additively. However, it is illuminating to work out the consequences of the simple model, and to compare these with the results of experiments, particularly on artificial selection. One can then ask what changes in the model are needed to explain the facts.

Phenotypic distributions

Figure 6.2 shows the distribution of some quantitative traits. Their common characteristic is that they are approximately normal, or Gaussian. This is what we expect on our model. If only a few loci are involved, we expect the phenotypic distribution to be skewed (Fig. 6.3A), unless allele frequencies happen to be 0.5, but as the number of loci increases, the skew disappears.
Phenotypic distributions

![Frequency distribution of four quantitative traits, with normal curves superimposed: A mouse, growth from 3 to 6 weeks of age; B mouse litter size; C Drosophila melanogaster, abdominal bristle number; D D. melanogaster, number of eye facets in the mutant Bar. (After Falconer 1981.)](image)

(Fig. 6.3B), even if allele frequencies are not 0.5, provided they are not very extreme. (Mathematically, this follows from the fact that the binomial distribution, \((p + q)^n\), tends to the normal as \(n\) increases.)

It is important to realize, however, that the prevalence of normal distributions
does not prove that our model is correct. A normal distribution is expected whenever a number of separate causes act additively: the causes could equally well be environmental. Thus in Fig. 6.1, of variation in an isogenic line, the causes cannot be genetic.

A second reservation is that phenotypic distributions are not always Gaussian. One reason is as follows. Imagine a population of geometrically similar organisms of different absolute size. If their heights are normally distributed, their weights cannot be, and vice versa, because weight \( \propto (\text{height})^3 \).

**Resemblances between relatives**

How similar, on the additive genetic model, do we expect relatives to be, when compared to two random members of the population? The most direct approach is to ask: what fraction of their genes are identical by descent? That is, what fraction are identical because they are copies of a gene in a relative. Remember that if we sample two genes at a locus randomly from the population, they may be identical, and they may not. What we are interested in, however, is the additional similarity arising from the genetic relationship.

We therefore imagine the genes of an individual as being of two kinds: those that are identical copies of genes present in a relative, and those that are a random sample of the genes in the population. This approach is applied in Figs 6.4 and 6.5, for autosomal genes in a diploid population. The results are summarized in Table 6.2.

Suppose, then, that we have pairs of measurements, \( z_1 \) and \( z_2 \), of some trait—say height—of pairs of relatives. The mean values are \( \bar{z}_1 \) and \( \bar{z}_2 \). Since we are assuming that all differences are genetic, and that genes act additively, there should be some measure of resemblance which, if our model is true, has the values shown in Table 6.2. The appropriate measure is the covariance, or, more precisely, the correlation coefficient. Thus
Phenotypic distributions

FIG. 6.4. Genes in common between parents and offspring. A and B are first sibs, and half-sibs. Offspring A received half her genes from father (F), and half from mother (M). B is a full sib of A, and has half her genes identical to those in A—one-quarter through F and one-quarter through M. C is a half-sib of A, with a different father, F', but the same mother. Only one-quarter of C's genes are identical to genes in A.

FIG. 6.5. Genes in common between cousins. A and B are first cousins. Consider A's genes. One-quarter come from grandfather (GF) and one-quarter from grandmother (GM). Of these, one-quarter were transmitted to B. Hence the fraction of B's genes that are shared with A is \( (1/4)^2 = 1/16 \).

\[
\text{Cov} (z_1, z_2) = \frac{1}{n} \sum (z_i - \bar{z}_1)(z_i - \bar{z}_2),
\]

and the correlation between \( z_1 \) and \( z_2 \) is

\[
r = \text{Cov} (z_1, z_2) / \sigma_1 \sigma_2,
\]

where \( \sigma^2 = \frac{1}{n} \sum (z_i - \bar{z}_1)^2 \) and \( \sigma^2 = \frac{1}{n} \sum (z_i - \bar{z}_2)^2 \).
<table>
<thead>
<tr>
<th>Parent–offspring</th>
<th>1/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full sibs</td>
<td>1/2</td>
</tr>
<tr>
<td>Half-sibs</td>
<td>1/4</td>
</tr>
<tr>
<td>First cousins</td>
<td>1/8</td>
</tr>
</tbody>
</table>

It is shown in Box 6.1 that the correlation between parent and offspring for our model is indeed one-half, corresponding to the fact that they have half their genes identical by descent. The same correspondence can be shown, with more difficulty, for other relationships.

It is helpful to have some intuitive feel for why the correlation coefficient is the right measure. This is done in Box 6.2, where it is shown that $r$ does indeed measure the causes that are common to the members of a pair, as a fraction of the total causes of variation.

Suppose that we find, in some population, that the correlation between the heights of sibs is indeed one-half. Does this prove that our model is correct, and in particular that all variance is caused by genes with additive effects? It does not, for two reasons:

1. In many species, sibs share an environment as well as genes. Sib correlations of approximately one-half are not uncommon for human traits, but by themselves they prove little, because the common causes may be environmental.

2. Genes that act non-additively may cause correlations between sibs, but not between parent and offspring. This is illustrated in Table 6.4, for a trait determined by a single overdominant locus, with phenotypes $aa = 0$, $Aa = 1$, and $AA = 0$, and allele frequencies of $a$ and $A$ of 0.5. There is a resemblance between sibs, because some families are all 0, and some are all 1. But there is no correlation between parent and offspring: $AA$, $Aa$, and $aa$ fathers have, on average, the same proportions of different kinds of offspring.

It follows that a sib correlation of 0.5 proves rather little. But a correlation of 0.5 between parent and offspring, if environmental correlations can be ruled out, would indicate that our model is close to the truth for that trait.

### The effects of inbreeding

What does our model predict as the result of continued inbreeding—that is, mating together close relatives? Suppose that, starting from an outbred population, we mate brother and sister in every generation. Consider a single locus, with two alleles, $a$ and $A$, segregating in the initial population. As we inbreed, sooner or later we will mate $AA \times AA$ or $aa \times aa$. Once that has happened, the
Box 6.1. The correlation between parent and offspring

In a diploid random-mating population, the value of some trait is determined by two alleles at a locus, with values $aa = 0$, $Aa = 1$, and $AA = 2$. The frequency of allele $a$ is $p$. Table 6.3 shows the mean values of sons from each type of father. In calculating the offspring from, say, an $Aa$ father, we note that the father contributes alleles $a$ or $A$ with probability $1/2$, and that, in either case, the mother contributes alleles $a$ and $A$ with probabilities $p$ and $q$, respectively.

### Table 6.3. The parent–offspring correlation at an additive locus

<table>
<thead>
<tr>
<th>Father Genotype</th>
<th>Frequency</th>
<th>Phenotype</th>
<th>Son $aa$</th>
<th>$Aa$</th>
<th>$AA$</th>
<th>Mean phenotype of sons</th>
</tr>
</thead>
<tbody>
<tr>
<td>$aa$</td>
<td>$p^2$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>$q$</td>
</tr>
<tr>
<td>$Aa$</td>
<td>$2pq$</td>
<td>1</td>
<td>$p/2$</td>
<td>$q/2$</td>
<td>$q/2$</td>
<td>$1/2 + q$</td>
</tr>
<tr>
<td>$AA$</td>
<td>$q^2$</td>
<td>2</td>
<td>$q$</td>
<td></td>
<td></td>
<td>$1 + q$</td>
</tr>
</tbody>
</table>

Now

$$
\sum(x - \bar{x})(y - \bar{y}) = \sum xy - \bar{x} \sum y - \bar{y} \sum x + n\bar{x}\bar{y} = \sum xy - n\bar{x}\bar{y}.
$$

For our model, $n = 1$ (since we are working with frequencies), $\bar{x} = \bar{y} = 2q$, and $\sigma_x^2 = \sigma_y^2 = 2pq$. From the table:

$$
\sum xy = 0 \times p^2q + 2pq(1 + q) + 2q^3(1 + q)
= pq + 2pq^2 + 2pq^3
= pq + 2q^2(1 + q)
= pq + 4q^2.
$$

Hence

$$
Cov(xy) = pq + 4q^2 - (2q)^2 = pq,
$$

and

$$
r = \frac{Cov(xy)}{\sigma_x \sigma_y} = \frac{pq}{pq/2pq} = 0.5.
$$

Hence the correlation between father and son (or, since we are considering autosomal genes, between parent and offspring of either sex) is one-half. If the trait is affected by genes at many loci, the value of $r$ is unaltered, provided that genes at different loci combine additively.

line will be genetically homozygous at that locus indefinitely, barring new mutation. If the initial population was segregating at many loci, an inbred line will become homozygous at successively more loci. The rate of this process is treated theoretically on p. 140. For the present, however, we can make a number of qualitative predictions from our model.

1. An inbred line will become phenotypically more uniform, until finally all members are identical. This, of course, does not happen, because inbreeding
Box 6.2 The correlation coefficient

Suppose that we have measurements, $z_1$ and $z_2$, on pairs of relatives, for example sibs. It is convenient to take these measures as departures from the mean values: hence the mean value of $z_1$ and of $z_2$ is zero. The correlation coefficient is then

$$r = \frac{\sum z_1 z_2}{(\sum z_1^2 \sum z_2^2)^{1/2}}. \quad (6.5)$$

The values of $z_1$ and $z_2$ are made up of two components, one of which is common to the two members of a pair, and the other of which takes independent values for the two members. That is

$$z_1 = x_1 + y_1; \quad z_2 = x_2 + y_2,$$

where $x_1$, $x_2$ are the common components, and $y_1$ and $y_2$ the independent components. Thus $x_1 = x_2$ for each pair. Then

$$\sum z_1 z_2 = \sum (x_1 + y_1)(x_2 + y_2) = \sum x_1^2 + \sum x_1 y_2 + \sum x_2 y_1 + \sum y_1 y_2.$$

Now the expected value of a term such as $\sum x_1 y_2$, consisting of the product of two independent variables, is zero. This is because the cases in which the two values have the same sign ($++$ or $--$) are exactly balanced by those in which they have opposite signs ($+-$ or $-+$). Hence

$$\sum z_1 z_2 = \sum x_1^2, \quad (6.6a)$$

and, by a similar calculation,

$$\sum z_1^2 = \sum x_1^2 + \sum y_1^2; \quad \sum z_2^2 = \sum x_2^2 + \sum y_2^2. \quad (6.6b)$$

In the case of pairs of relatives, the values of the independent components of variance, $\sum y_1^2$ and $\sum y_2^2$, will be equal. Hence, substituting from Equations 6.6a and 6.6b into Equation 6.5, we have

$$r = \frac{\sum x_1^2}{(\sum x_1^2 + \sum y_1^2)}. \quad (6.7)$$

We can therefore regard $r$ as a measure of the variance that is common to the members of a pair, as a fraction of the total variance. For our additive genetic model, the value of $r$ is equal to the fraction of genes in the two members of a pair that are identical by descent.

does not eliminate environmental variance. However, if genetic and environmental effects are additive and independent, so that $V = V_G + V_E$, we would expect phenotypic variance to decline, as $V_G$ tends to zero. This expectation is often not realized: it is common to find that inbred lines are more variable than the outbred populations from which they were derived. This is evidence that inbreeding reduces the capacity of organisms to
regulate during development: inbreeding reduces canalization, or developmental homeostasis.

2. An inbred line will become genetically uniform, and will no longer respond to artificial selection. This prediction is born out by experiment. In vertebrates, a second proof of the genetic uniformity of inbred lines is possible: skin can be grafted between members of a line.

Inbreeding has one other effect that is not predicted by the additive model. It is usually accompanied by a massive decline in fertility, viability, and other fitness components. This is illustrated in Fig. 6.6. One reason for inbreeding depression is that the line becomes homozygous for deleterious recessives that were present in the initial population. A second possible reason is that there are some loci at which the heterozygote is fitter than either homozygote (see p. 65); if so, inbreeding will inevitably lead to a decline in vigour. There has been considerable argument about whether this second effect is important: the causes of inbreeding depression are discussed further in Box 6.3.

### The effects of directional selection

Figure 6.9 defines some of the terms used to describe directional selection:

- The selection differential, $S$, is the difference between the mean value of the selected parents, and that of the population as a whole.

- The response, $R$, is the difference between the mean value of the offspring generation, and that of the population in the previous generation.

- The intensity of selection, $l$, is $S/\sigma_p$, where $\sigma_p$ is the phenotypic standard deviation of the population before selection.
Box 6.3 The causes of inbreeding depression

The phenomena to be explained are as follows. Inbred lines derived from naturally outbreeding populations show a decline in viability, fertility, and growth rate. When inbred lines are crossed, the $F_1$ hybrids are usually as vigorous as the members of the original outbred population: that is, they show hybrid vigour (see Fig. 6.7 for an example).

There are two possible reasons for inbreeding depression:

1. **True overdominance**: that is, there are loci at which the heterozygote is fitter than either homozygote.

2. **Associative overdominance**: Different inbred lines become homozygous for different deleterious recessive genes. For example, one line might have the genotype $m_1+/m_1+$, and another $+/m_2/m_2$. The $F_1$ between them would be $m_1++/m_2$, and would be of high fitness, because the deleterious genes $m_1$ and $m_2$ are recessive.

It is certain that some part of the decline in fitness is caused by deleterious recessive genes. As inbreeding continues, however, the more serious recessives (e.g. lethals) will be eliminated by selection. Lines which do not die out recover somewhat in vigour, as shown in Fig. 6.6. They do not recover fully, because, by chance, some deleterious alleles will become fixed, and once this happens only back mutation can remove them.

How can we decide whether truly overdominant loci are also important?
Phenotypic distributions

FIG. 6.7: An example of hybrid vigour. Survival curves for two inbred lines, B and K, of Drosophila subobscura, and the F₁ hybrids, B/K and K/B, between them (sexes combined). (From Clarke and Maynard Smith 1965.)

There are two possible methods. One is to identify the particular loci involved. Box 4.3 presented evidence that some enzyme loci that are polymorphic in natural populations are truly overdominant. But we have no evidence that such loci are important in inbreeding depression.

The second is to use the methods of quantitative genetics. Suppose that inbreeding depression is caused by homozygosity for regions of chromosome. I leave open for the moment whether this is true or associative overdominance. Let us call the chromosome regions a and A, without implying that these are single genes. The phenotypes are aa = 0, Aa = 1, and AA = 0; that is, homozygotes are of low vigour.

Two inbred lines have the genotypes aa and AA. Figure 6.8 shows the genotypes and phenotypes of the F₁ and F₂ hybrids, and also of progeny obtained by backcrossing F₂ individuals to the parental lines. It has been assumed in the figure that chromosome regions a and A behave as units; that is, if the cause is associative overdominance, the loci would have to be fairly tightly linked.

Notice the following facts:

1. Members of the F₂ differ in phenotype.
FIG. 6.8. \( F_1 \), \( F_2 \), and backcross generations between two inbred lines, assuming that the phenotypes are \( aa = 0 \), \( Aa = 1 \), and \( AA = 2 \). Mean phenotypes are shown in parentheses. Note that the three \( F_2 \) genotypes have different phenotypes, but the average phenotypes of their backcross progeny are the same.

2. The offspring of different \( F_2 \) individuals, averaged over the two backcrosses, always have the same mean values. However, if \( A \) were dominant to \( a \), there would be differences between the mean values of the offspring of different \( F_2 \) individuals.

Therefore, by comparing these two kinds of variance—between \( F_2 \) individuals, and between the mean values of their backcross offspring, we can get a measure of how much overdominance there is among the \( F_2 \) (the exact method of analysing the variance is described by Bulmer 1980).

Moll et al. (1964) used this method to analyse crosses between inbred lines of maize. They found appreciable overdominance, but this could be from either of the two causes. They therefore interbred the \( F_1 \) for a number of generations, and repeated the analysis on the \( F_2 \). Now if the cause of overdominance is the presence of repulsion linkages between deleterious recessives \((m_1 + \text{ and } + m_2)\), this procedure, by allowing recombination and hence reducing linkage disequilibrium, should destroy the overdominance. This is in fact what they found. They concluded that, in maize, hybrid vigour can be accounted for by associative overdominance, and that there is no reason to assume the presence of truly overdominant loci.

If all or most of inbreeding depression is caused by homozygosity for deleterious recessives, it should be possible to produce an inbred line that is as vigorous as an outbred population. Thus different lines will be homozygous for different recessives. By crossing two lines, and then again inbreeding, one
should obtain a new line with fewer deleterious genes. Repeating this process should ultimately produce inbred lines of high fitness. In contrast, if true overdominance is widespread, no inbred line can be of high vigour. In practice, it has proved difficult to obtain vigorous inbred lines. However, there are natural experiments. Many plant species show a high frequency of self-fertilization, in nature, but almost always there is occasional outcrossing. This provides the ideal breeding system for studying many different homozygous genotypes and eliminating the less fit.

Do naturally selfing plants show hybrid vigour when they are crossed? The degree of hybrid vigour is certainly not as great as it is when crossing inbred lines derived from natural outbreeders, but there is evidence for a small degree of hybrid vigour. For example, the poppy *Papaver dubium* shows 5% percent selfing in the wild. Cole et al. (1976) derived a number of selfed lines of *P. dubium*. When they crossed lines derived from the same wild population, they found that mean capsule number was usually greater in the hybrids than in either line. However, this does not prove that overdominant loci are important. Mutation is a continuing process. Inbred lines will carry mutations for longer, and it is likely that different lines will carry different mutations, and will display different hybrid vigour.

How do we expect \( R \) to be related to \( S \)? For our model, the answer is simple: we expect \( R = S \). The reason is as follows. At each locus, the offspring generation has the same alleles, in the same frequencies, as the selected parent. Since all differences are genetic, and genes act additively, the mean offspring phenotype equals the mean parental phenotype: that is, \( R = S \). In practice, as will be discussed in Section 6.4, \( R < S \): the response to selection is not as great as the selection differential.

We can now summarize some of the predictions of the additive genetic model:

1. Phenotypic distributions will be Gaussian: the fact that many actual distributions are approximately Gaussian is therefore consistent with the model, but it is not strong confirmation of it.
2. The phenotypic correlations between relatives will be equal to the proportion of genes identical by descent.
3. Inbreeding will produce a population that is genetically and phenotypically uniform.
4. The response to selection, $R$, is equal to the selection differential, $S$.

We have already met some facts that do not agree with these predictions. In particular:

1. Inbred lines are phenotypically variable: however, the results of inbreeding do conform to prediction in that selection on an inbred line is usually ineffective.
2. Inbred lines usually show a loss of viability and fertility.
3. The response to selection is usually less than the selection differential.

The next section describes a more realistic model, allowing for environmental causes of variation, and for non-additive effects, which is able to account for these discrepancies.

A more realistic model

The most obvious fact that has been omitted from the model of Section 6.2 is that many differences are environmentally caused. I now introduce environmental variance, but for the time being I retain the additive assumption: that is, environmental and genetic factors act additively, as in Table 6.1A, and not as in
Table 6.1B, and genes act additively, as assumed in the last section. We can then regard the phenotype, $Y$, as the sum of a genetic and environmental component,

$$Y = G + E,$$  \quad (6.8)

and, if genotype and environment are independent, the total variance of $Y$ is the sum of the genetic and environmental variances,

$$V = V_G + V_E.$$

Clearly, the introduction of environmental variance can explain the fact that inbred lines are phenotypically variable: $V_G$ goes to zero, but $V_E$ remains. However, with the additive model we still cannot explain why, in some cases, $V$ is actually greater in an inbred line than in the outbred population from which it was derived. This requires non-additive effects, or gene–environment interaction: the effects of environmental factors on the phenotype are greater on an inbred than on an outbred genetic background.

**Box 6.4. The response to selection for the additive model**

In reading this Box, it will be helpful to remember the argument of Box 6.2, which was that a correlation coefficient measures the variance that is common to the members of a pair, as a fraction of the total variance.

For the additive model, $Y = G + E$, where $Y$ is the phenotype of an individual, and $G$ and $E$ are the genetic and environmental contributions, respectively. It is convenient to measure $P$ as a departure from the mean value of the population, the mean values of $G$ and $E$ in the population are also zero. If we select a set of parents of mean phenotype $Y_p$, and obtain offspring of mean phenotype $Y_o$, then the selection differential $S = Y_o - Y_p$.

The first point to establish is that, in any set of individuals, the mean value of $G$ depends only on the allele frequencies, and not on the genotype frequencies. Thus:

<table>
<thead>
<tr>
<th>Genotype at a locus</th>
<th>aa</th>
<th>aA</th>
<th>AA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution to phenotype</td>
<td>0</td>
<td>d</td>
<td>2d</td>
</tr>
<tr>
<td>Frequency</td>
<td>$P$</td>
<td>$Q$</td>
<td>$R$</td>
</tr>
</tbody>
</table>

Then the contribution of the locus to $G$ is $dQ + 2dR = 2dp$, where $p$ is the frequency of allele $A$. As the contribution depends on $p$ but not separately on the genotype frequencies; if loci combine additively (that is, no epistasis), then $G$ depends only on the gene frequencies.

The next point is this: if a set of individuals breed together and produce a new generation, then the mean value of $G$ among the offspring equals the mean
value among the parents. This follows from the fact that the gene frequencies among the offspring equal those among their parents.

We are now in a position to tackle the main problem. If we select a set of parents with mean phenotype $Y_p$, what will be the mean phenotype $Y_o$ of their offspring?

We have:

$$Y_o = Y_p + Z$$

and

$$Y = Y_p - Z$$

The second equation holds because the offspring have the same genetic contribution as their parents that are exposed to a typical range of environments with zero mean. The $Z$ is a random variable whose expected value is the expected value of $G$, given that we know $Y_p$. That is, we need to know the value of $Y_p$ in the equation

$$G = Y_p - Z$$

in order to find $E(Y)$. From statistical theory

$$E(Y_p) = Y_f Y_p$$

Here $E(Y_p)$ is the expected value of $Y_p$.

The relation between response to selection and selection differential is given by

$$R = h^2 S,$$  \hspace{1cm} (6.10)

where

$$h^2 = V_G / (V_G + V_e).$$  \hspace{1cm} (6.11)

In interpreting this equation, it is important that the derivation in Box 6.4
assumes additivity, and in particular that genes act additively. We have seen (Table 6.4) that there can be genetic variance, but no correlation between parent and offspring, and hence no response to selection. It is only the additive effects of genes that contribute to the response to selection. We should therefore rewrite Equation 6.11 as

$$h^2 = \frac{V_A}{V},$$

(6.12)

where $V_A$ is the variance due to the additive effects of genes, and $V$ the total phenotypic variance. So defined, $h^2$ is the heritability, or, more precisely, the narrow-sense heritability. In contrast, the broad-sense heritability is defined as $V_G/V$, where $V_G$ is the total genetic variance. Since only the additive effects of genes contribute to the response to selection, it is the narrow-sense heritability that should be used in the equation $R = h^2 S$.

How are we to measure $h^2$? The simplest way is to carry out a single generation of selection, and measure $R$ and $S$. Then $h^2 = R/S$. A heritability measured in this way is called a realized heritability.

An alternative is to measure $V_A$ from the resemblance between relatives: this method is described in Box 6.5.

One last point should be made about estimates of heritability. To say, for example, that $h^2$ for size in Drosophila melanogaster is 0.4 cannot be a universal
complete overdominance ($aa = AA = 0$, $Aa = 1$), there is a correlation between sibs, but not between parent and offspring. If there is no parent–offspring correlation, there can be no response to selection.

2. *Epistatic interactions between loci.* To illustrate this, consider a haploid organism with two equally frequent alleles at each of two loci. Suppose that two of the genotypes, $ab$ and $AB$, are selected, and the other two, $aB$ and $Ab$, are discarded. It is shown in Table 6.5 that, after the first generation, there is no response to selection. Yet there is genetic variation, and sibs do resemble one another. It is worth noting, however, that the equilibrium illustrated in the table is unstable (see Problem 9).

3. *Gene–environment interaction.* This is the phenomenon illustrated in Table 6.1B.

<table>
<thead>
<tr>
<th>Table 6.5. Selection with epistasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenotype</td>
</tr>
<tr>
<td>Zygote frequencies, generation $n$</td>
</tr>
<tr>
<td>Adult frequencies after selection</td>
</tr>
<tr>
<td>Zygote frequencies, generation $n+1$</td>
</tr>
<tr>
<td>Adult frequencies after selection</td>
</tr>
<tr>
<td>Zygote frequencies, generation $n+2$</td>
</tr>
</tbody>
</table>

Zygote frequencies are derived from adult frequencies as follows:

<table>
<thead>
<tr>
<th>Mating</th>
<th>Frequency</th>
<th>Genotype of offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ab \times ab$</td>
<td>$1/4$</td>
<td>$ab$ 0 0 0</td>
</tr>
<tr>
<td>$ab \times AB$</td>
<td>$1/2$</td>
<td>$aB$ 1/8 1/8 1/8 1/8</td>
</tr>
<tr>
<td>$AB \times ab$</td>
<td>$1/4$</td>
<td>$Ab$ 0 0 0 1/2</td>
</tr>
<tr>
<td>$AB \times AB$</td>
<td>$3/8$</td>
<td>$AA$ 1/8 1/8 3/8</td>
</tr>
</tbody>
</table>

Experiments in artificial selection

Figure 6.11 shows some of the results of an artificial selection experiment on the number of abdominal bristles in *Drosophila*: the results are typical of many such experiments. The following comparisons can be made with the predictions of the simple model:

1. The population did respond to selection in both directions. However, the response is asymmetric, being greater in the upwards-selected lines. Realized heritabilities in such experiments are less than one, because of environmental variance. Some values are given in Table 6.6: these are discussed further on p. 118.
truth. At best, it is true of a particular population in a particular range of environments. Thus $h^2$ measures the additive genetic variance as a fraction of the total variance. Any change that reduces the genetic variance, or increases the environmental variance, will reduce $h^2$.

Before turning to the empirical data, it is useful to review some of the kinds of genetic variation that does not contribute to a selective response. There are three main categories:

1. Dominance interactions between alleles. It was shown in Table 6.4 that, with
2. The population reached a selection limit, beyond which further progress was difficult or impossible. Such limits are typical in laboratory experiments on artificial selection. They arise because the additive genetic variance present in the initial population has been fixed. The length of time taken to reach a limit, and the difference between the initial and final populations, depend on the number of loci involved; the matter is discussed further in Box 6.6.

Table 6.6. Approximate values of realized heritability for traits in *Drosophila*

<table>
<thead>
<tr>
<th>Trait</th>
<th>$h^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>D. melanogaster</em></td>
<td></td>
</tr>
<tr>
<td>Abdominal bristle number</td>
<td>0.5</td>
</tr>
<tr>
<td>Body size (thorax length)</td>
<td>0.4</td>
</tr>
<tr>
<td>Ovary size</td>
<td>0.3</td>
</tr>
<tr>
<td>Egg production</td>
<td>0.2</td>
</tr>
<tr>
<td><em>D. subobscura</em></td>
<td></td>
</tr>
<tr>
<td>Development rate (days from egg to adult)</td>
<td>0.2</td>
</tr>
<tr>
<td>Slow-selected</td>
<td></td>
</tr>
<tr>
<td>Fast-selected</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Box 6.6 The number of loci involved in the response to selection

Suppose that, in the initial population, there are \( n \) loci affecting some trait, each with two additive alleles, with frequencies \( p \) and \( q \). The difference between the two homozygotes is \( 2d \). Then the difference, \( D \), between the mean values of the trait in up- and down-selected lines, when the selection limit is reached, is 2\( nd \). The additive genetic variance in the initial population is \( 2npqd \). Both \( D \) and \( V_A \) can be measured. Then

\[
D^2/V_A = 4n^2d^2/2npqd^2, \\
= 2n/pq.
\]

Hence

\[
n = \frac{pqD^2}{2V_A}.
\]

It was suggested in the main text that the initial frequencies of the relevant loci are intermediate, essentially because there is no evidence of a large increase of \( V_A \) as selection proceeds. If \( p = q = 1/2 \), then

\[
n = \frac{1}{8V_A} D^2. \quad (6.13)
\]

Using this formula, Falconer estimates that the number of loci involved in some selection experiments in mice and Drosophila is in the range 30–100.

The largest potential source of error in such an estimate lies in the assumption that \( p = q = 1/2 \). If \( p = 0.01 \), then \( n = D^2/200V_A \). Falconer’s estimates would then lie in the range 1–5 loci. This is almost certainly an underestimate, but the value of 30–100 is probably an overestimate.

3. After the selection limit was reached, populations in which selection was relaxed did not tend to return towards the original state, suggesting that there was little non-additive genetic variance for the trait. This is a little unusual in many experiments there is some evidence for non-additive genetic effects, both in the presence of sib correlations at the selection limit, and in the tendency of selected populations to return towards their original state when selection is relaxed.

4. Flies in the final populations were of lowered fertility. Such correlated responses to selection are common. They may be caused by pleiotropic effects of the selected alleles, or by linkage disequilibrium between the selected alleles and loci affecting other traits: remember that these experiments are usually carried out on rather small populations, so that linkage disequilibrium due to chance is bound to be present.

One final question is crucial if we wish to extend these conclusions from artificial selection to evolution. How far does the existence of a selection limit
depend on the fact that these experiments involve intense selection on a single trait in a small population? In these experiments, the response depends almost entirely on genetic variance present at the start. The number of new mutations, occurring after the start of the experiment, will be proportional to the number of generations, and to the number of parents in each generation: if both these are small, new mutations can play little part in the response.

Experiments in which selection was practised for many generations on a larger population suggest that selection limits may indeed be an artefact of small population size: Fig. 6.12 shows the results of such an experiment. In each line, 50 parents of each sex were selected, out of a total of 500 flies scored for abdominal bristle number in each generation. Selection was continued for 86–89 generations. Averaged over six replicate lines, the initial number was 8.2, and the final number was 33.9 bristles, a fourfold increase. The increase was 16 times the phenotypic standard deviation in the initial population. In quantitative terms, this is a greater response than the increase in brain size between *Australopithecus* and ourselves, of which we are so proud. In evolutionary terms, however, 50 pairs is a small population, and a selection of 20 per cent intense. This is reflected in some details of the response. Most lines showed

**FIG. 6.12.** Response to long-term selection for abdominal bristle number, using 50 parents of each sex in each generation; six replicate lines were run (after Yoo 1980).
Quantitative variation and fitness

The equation $R = h^2 S$ relates the response to selection on some trait to the selection differential. Suppose, now, that the trait under consideration is not wing length or bristle number, but fitness—that is, expected number of offspring. A population, before selection, consists of a number of genotypes, $g_1, g_2, \ldots, g_i, \ldots$. Let the frequency of $g_i$ be $p_i$, and its fitness be $w_i$. Then we can define the mean fitness as

$$\bar{w} = \sum p_i w_i. \quad (6.14)$$

After selection has operated, the frequency of genotype $g_i$ is $p_i w_i / \bar{w}$, and hence the mean fitness of the selected parents is

$$\bar{w}' = \sum p_i w_i^2 / \bar{w}.$$ 

Hence the selection differential on fitness is

$$S = \bar{w}' - \bar{w} = \sum p_i w_i (w_i - \bar{w}) / \bar{w}.$$ 

We want now to show that $S$ is equal to $V_w$, the variance of fitness before selection. Thus

$$V_w = \sum p_i (w_i - \bar{w})^2 = \sum p_i w_i (w_i - \bar{w}) - \sum p_i \bar{w} (w_i - \bar{w}),$$

and since the second term is zero,

$$V_w = \sum p_i w_i (w_i - \bar{w}),$$

and since, in a density-regulated population, $\bar{w} = 1$, we have $S = V_w$. Remem-
bering that \( h^2 = V_a / V_e \), and that \( R \) is the change in the selected trait in one generation, Equation 6.10 becomes

\[
\Delta \tilde{w} = V_a,
\]

(6.15)

where \( \Delta \tilde{w} \) is the increase in mean fitness in one generation.

This is a version of Fisher’s ‘fundamental theorem of natural selection’. His formulation was ‘the rate of increase of fitness of any organism at any time is equal to its genetic variance at that time’. Clearly, by ‘genetic variance’ he meant additive genetic variance. Thus, if the genetic variance of a population was due entirely to genes with heterotic effect (see p. 100), and the population was at a selective equilibrium, then both \( \Delta \tilde{w} \) and \( V_a \) would be zero, so Equation 6.15 would be true, but the total genetic variance of fitness, \( V_e \), would not be zero.

Fisher thought that his theorem could play the same role in biology as is played by the second law of thermodynamics in physics, by placing an arrow on time. Since a variance cannot be negative, Equation 6.15 implies that \( \tilde{w} \) can only increase, as entropy increases. There has been much subsequent debate about the theorem. My own view is that it cannot play an important role in biology. If it were true, it should be the case that natural selection necessarily increases the mean fitness of a population in some meaningful sense: for example, that it can maintain a larger population size, or is better able to survive a change in the environment, or competition from other species. Unfortunately, none of these conclusions follow. Consider, for example, the following plausible case. Selection within a species favours the larger individuals, because they are better able to defeat others in competition for scarce resources. The result is a steady increase in size, beyond the level that would be optimal in the absence of intraspecific competition. Hence the species becomes rarer, and less able to survive competition from other species. This illustrates one reason why Equation 6.15 can lead to misleading conclusions. It is based on the assumption that the fitness of a genotype is constant, and independent of what other genotypes are present. This is often not the case.

There is, however, one implication of Equation 6.15 that is illuminating. In most populations, the additive genetic variance of fitness will be small, because such variance will be used up by natural selection, just as artificial selection exhausts the additive variance for the selected trait. This conclusion is confirmed by the data in Table 6.6, showing that \( h^2 \) tends to be small for traits directly contributing to fitness. Three points need to be made:

1. In some cases, there was substantial non-additive genetic variance for fitness-related traits.
2. It would be wrong to conclude that the heritability of fitness will be zero, both because recurrent deleterious mutations will cause some heritable
fitness differences, and because, in a changing environment, populations are not in equilibrium under selection.

3. Even if the heritability of fitness is small, there may be substantial heritability for particular components of fitness, if different components are negatively correlated. In *Drosophila*, for example, Rose and Charlesworth (1980) found that those genotypes that lay eggs rapidly when young tend to be short lived, and vice versa. This is not surprising, because there is evidence that, in females of the same genotype, laying eggs shortens life.

The maintenance of genetic variance for quantitative traits

Why is there polygenic variation in natural populations? Ultimately, the origin is mutation, but why has not natural selection eliminated the less fit variants? From the arguments of Chapter 4, we can see that there are two possible answers:

1. There is a balance between mutation and selection.
2. There is a balance of selective forces (heterosis, frequency-dependence, selection for different types in different places or at different times).

Until relatively recently, the first of these possibilities was not taken very seriously as an explanation of polygenic variability. The arguments on page 55 led us to think that, if a mutant allele reduces fitness, even slightly, then its frequency in a natural population would be very low—essentially because mutation rates are very low. Now there are good reasons for thinking that the quantitative variation discussed in this chapter cannot be caused by loci with one common and one (or more) very rare alleles. If it were so, the genetic variance of a population would increase when it was exposed to directional selection (see Box 6.6, and Problem 10), and this does not usually happen. Hence, it was generally assumed that variation was maintained by a balance of selective forces.

This assumption has recently been challenged by Lande (1975), who has argued that quantitative variation is maintained by recurrent mutation. He suggests that most populations, most of the time, are under normalizing selection. If so, and if mutations at many loci can affect each quantitative trait, then appropriate genetic variance could be maintained, without need to invoke opposed selective forces.

This claim is still controversial. Arguments about it tend to involve complex mathematics, and extensive computer simulations. However, it is possible to grasp what the argument is about. Essentially, Lande's claim rests on one idea, and one observation. The idea is as follows. If a population is under normalizing selection for a polygenic trait, in the absence of mutation, it will ultimately lose most or all of its genetic variance, but the loss of variance will be slow. The
reason is illustrated in Fig. 6.13. If, then, normalizing selection is slow to eliminate genetic variance, it is more reasonable that mutation should maintain it.

The fact concerns the rate at which new heritable variation is generated by mutation. To measure this, one first produces a genetically homogeneous population, by inbreeding or chromosome manipulation, and then watches to see how rapidly heritable variation reappears. Such experiments have been performed on *Drosophila*, maize, and mice, and on several traits in each species. All give results on the order of

\[ V_m = \frac{V_e}{1000}, \]  

(6.16)

where \( V_m \) is the new heritable variation arising by mutation in one generation, and \( V_e \) is the environmental variance. Since, for most traits, \( h^2 \) is on the order of 0.5, this is equivalent to saying that 1/1000th part of the genetic variance is regenerated by mutation each generation.

It is, I think, too early to say whether polygenic variation is in fact maintained by a balance between mutation and selection. However, the empirical results (Equation 6.16) are somewhat paradoxical. The nature of the paradox, and

---

**FIG. 6.13.** The effect of normalizing selection on the gene frequency at a locus. It is assumed that the population phenotype is normally distributed about the optimum value, OPT. The distributions aa, Aa, and AA are the phenotypic distributions of individuals with those genotypes at a particular locus. For the particular case shown, \( p(a) < 0.5 \) and \( p(A) > 0.5 \): normalizing selection will further reduce \( p(a) \), until \( A \) is fixed. However, the mean of \( Aa \) individuals is close to the optimum, so the rate of approach to genetic homozygosity will be slow.
The maintenance of genetic variance for quantitative traits

some possible resolutions, are discussed in Box 6.7. The matter is discussed further by Turelli (1986).

Box 6.7 Is polygenic variation maintained by a balance between mutation and normalizing selection?

Suppose that there are $n$ loci affecting a particular trait. Let $u$ be the per locus, per generation mutation rate, and $m$ the effect of a single mutation on the phenotype (of course, not all mutations will have the same effect: $m$ is the root mean square of the value). Then, on the additive model, in a diploid population, the new genetic variance generated per generation is

$$2 \sum um^2.$$

If, in a typical outbred population, $v$ is the variance contributed by one locus, and if $h^2 = 0.5$, then Equation 6.16 implies

$$2 \sum um^2 = 10^{-3} \sum v.$$  \hspace{1cm} (6.17)

In Section 4.2 it was concluded that the mutation rate per gene, per generation is on the order of $10^{-5}$. If so, $v = m^2 / 50$. But if the variance was caused by two or more alleles of approximately equal frequency, $v$ would be of the same order as $m^2$. This suggests that the standing genetic variance at a locus is caused by one common allele and one or more rare alleles. But, as explained in the main text, there are good reasons for thinking that polygenic variation is not caused by rare alleles: if it were, directional selection would increase genetic variance.

We are therefore faced with a paradox. There seem to be two possible ways out:

1. The polygenic mutation rate is much higher than $10^{-5}$ per locus, per generation. This could be so. Thus the mutation rates estimated in Section 4.2 were for genes coding for proteins. As we shall see in Chapter 11, there is a great deal of DNA in the eukaryotic genome that is not translated. It is possible that changes in this DNA, while not altering the amino-acid sequence of any protein, may alter the rates or times at which proteins are synthesized. If so, such mutations could affect quantitative traits. This suggestion is highly speculative, but it cannot at present be ruled out.

2. There is something misleading about the empirical result summarized in Equation 6.16. It could be that the new heritable variability on which this estimate is based is not typical of the standing variability found in natural populations. Thus the new variability may involve mutations of relatively large effect (large $m$), whereas the standing variability involves allele differences of smaller effect.
Further reading


Problems

(Problems 3 and 4 are from J. F. Crow 1986.)

1. Two alleles are segregating at each of six loci in a random-mating diploid population. At each locus, the phenotypic values of the genotypes are \(-/\) = 0, \(-/+\) = 1, \(+/+\) = 2. The broad-sense heritability is one. Genes are additive between loci, so that the phenotype of an individual that is \(-/-\) at all six loci is 0, and of one that is \(+/+\) at all loci is 12. The frequencies of the \(+\) allele at the six loci are 0.2, 0.3, 0.4, 0.6, 0.7, and 0.8. (a) What proportion of the population has phenotype 12? (b) What is the phenotypic variance of the population?

2. A population has genetic variance as described in Problem 1. There is also environmental variance. The total phenotypic variance is 6. Parents are selected whose phenotype is 1 unit above the population mean. What is the response to selection?

3. What is the heritability of sex?

4. Would you believe someone who told you that the correlation in intelligence between half-sibs is 0.3?

5. What is the coefficient of relatedness of a niece to her aunt?

6. In a population of *Drosophila*, the mean abdominal bristle number is 18. The correlation between half-sibs is 0.1. A set of parents are selected with a mean number of 21. What is the expected bristle number of their offspring?

7. A trait is determined by two equally frequent alleles at a locus. The phenotypes are \(aa\) = 0, \(aA\) = \(AA\) = 2. (a) What is the total variance, the additive genetic variance (that is, the variance if the phenotype of \(aA\) is \(t\)), and the narrow-sense heritability? (b)* There is selection, with fitnesses \(aa\) = 1, \(aA\) = \(AA\) = \(1 + s\). Calculate the selection differential, and the expected value of the realized heritability.

8.* A population of *Drosophila* has a mean abdominal bristle number of 24, with a standard deviation of 3. The realized heritability is 0.3. Selection limits were reached at 32 bristles in the up line, and 16 bristles in the low line. Assuming that the variance in the original population was caused by loci with two alleles at approximately equal frequencies, estimate the number of loci concerned.

9. Suppose that, in the example of Table 6.5, the adult frequencies after selection were 0.6 \(ab\), 0.4 \(AB\). What will be the frequencies in the next generation?

10. The genetic variance of a population is caused by genes with additive effects. At
half the relevant loci the + allele has a frequency of 0.99 and the − allele of 0.01, and at the remaining loci the frequencies of the + and − alleles are reversed. If the magnitude of the effect of an allele substitution is the same at all loci, would you expect the genetic variance to increase or to decrease under directional selection, and by how much?

Computer projects

Usually, the simulation of polygenic inheritance requires a larger computer, and more sophisticated programming, than I have assumed in other chapters. There are two methods of computer analysis. The first—the Monte Carlo method—assumes a finite population. The genotypes of all the individuals in one generation are stored in an array. Two parents are chosen randomly, and one or more offspring are generated according to Mendelian laws and placed in an array holding the next generation, the process being repeated until the required number of offspring have been produced. The method has the advantage of realism, but only rather small populations can be simulated in a reasonable time. The alternative is to assume an infinite population, and calculate the genotype frequencies in the next generation deterministically. This can take a lot of computer time if the number of loci is large. Thus suppose there are six linked loci, with two alleles per locus. There are 64 gamete types whose frequencies must be stored in an array. To produce the next generation, assuming random mating (matters can become horrendous if we do not), we consider in turn the 2080 diploid genotypes (why 2080?) and calculate what proportion each one produces of the 64 gamete types. That takes time. If we add one locus, that would increase the time by a factor of 8. Things should get better when parallel-processing computers are available.

Because of these difficulties, the following project is suitable only for those with some computing experience:

A diploid organism has up to (say) six linked loci, with two alleles per locus. Write a procedure which, given the genotype of an individual, calculates the frequencies of all the gamete types produced, and adds those frequencies to an array holding the frequencies of the 64 possible gamete types. Assume a cross-over frequency of \( r \) between neighbouring loci. The procedure should allow for single and multiple crossing over. Assume no interference. Write the procedure in a language that permits you to store a gamete as a binary number (with 0 and 1 specifying the alleles), that will convert binary into decimal numbers and vice versa, and that allows you to use operators AND and OR on single elements of a binary number. (A procedure of this kind is the guts of any program that simulates quantitative genetics.)
3. \( u = p_{sh} \cdot p = 10/2 \times 94.075 \). Fitness of heterozygotes = \( 27/108 \times 487/582 = 0.21 \). Hence \( hs = 0.79 \), and \( u = 4.2 \times 10^{-5} \).

4. (a) \( p = \frac{1}{s} (u/s) \), or \( u = p^2 s \) where \( p^2 = 1/10 \, 000 \) and \( s = 1 \). Hence \( u = 1/10 \, 000 \).
	(b) If the relative fitnesses are \( 1:1 - \delta \), then from Equation 4.5 \( p = \delta/(1 + \delta) \). Since \( p = 1/100 \), \( \delta = 1/99 \).

6. \( p(A) = 0.9091; \bar{w} = 0.9545 \), so load = 0.0455.
7. Fitness of homozygote = 0.889; \( \bar{w} = 0.90 \), so load = 0.1.

8. At equilibrium, \( 0.75 + p = 1.5 - P \), or \( P = 0.375 \). If \( p \) is the frequency of \( A \), then \( (1 - p)^2 = 0.375 \), or \( p = 0.388 \). The equilibrium is stable, because dark is fitter than pale when dark is rare, and pale is fitter than dark when pale is rare. The genetic load at equilibrium is zero, because all three genotypes have the same fitness.

9. Relative to \( M \) flies, the viability of \( +/+ \) is \( 467/2 \times 201 = 1.1617 \), and of \( +/1+ \) is \( 376/2 \times 197 = 0.9543 \). Hence the relative viability of \( +/+1 \) is 0.821.

10. If there are no cell deaths, one cell gives rise to \( 10^8 \) cells by \( 10^8 - 1 \) cell divisions. If the back mutation rate = \( u \), the probability of no back mutations in a tube is \( (1 - u)^{10^8} = \exp(-10^8 u) = 0.72 \), or \( u = 3.28 \times 10^{-6} \).

Chapter 5

1. \(-0.02\).
2. \(-0.02 \times -0.9^4 = -0.0131\).

3. Yes. Writing the frequencies of the four phenotypes as \( p(AB) \), \( p(AB) \), \( p(ab) \), and \( p(ab) \), we expect \( p(AB) \cdot p(ab) = p(AB) \cdot p(ab) \) if there is linkage equilibrium.

4. (a) Zero. (b) Negative. Let the initial frequencies of the haplotypes be \( p(AB) \), \( p(AB) \), \( p(ab) \), and \( p(ab) \), where \( p(AB) \cdot p(ab) = p(AB) \cdot p(ab) \). In case (a) \( \Delta \) after one generation is \( p(AB) \cdot p(ab) \cdot (1 + s)(1 + t) -(1 + s)(1 + t)\) = 0. In case (b) \( \Delta = p(AB) \cdot p(ab) \cdot (1 + s + t -(1 + s)(1 + t) = -s \cdot p(AB) \cdot p(ab) \). It is an important result, due to Felsenstein, that if a population starts in linkage equilibrium, and fitnesses are multiplicative (case a), it remains in linkage equilibrium. But if the population is exposed to directional selection, and fitnesses increase less steeply than in the multiplicative case, then negative linkage disequilibrium is produced.

Chapter 6

1. (a) \( 6.5 \times 10^{-5} \); (b) 2.44.
2. Heritability = 2.44/6, so response \( = 1 \times h^2 = 0.407 \).

3. Sex is genetically determined, but there is no correlation between the sex of a child and a parent. So the broad-sense heritability \( = 1 \), and the narrow-sense heritability \( = 0 \).

4. If the half-sibs were raised apart, this implies a heritability of \( 4 \times 0.3 = 1.2 \), which is not possible. But if they are raised together, or in similar environments, a value of 0.3 is quite plausible.
5. $1/4$.

6. $h^2 = 0.4$, and $S = 3$, so expected value of offspring is $18 + 0.4 \times 3 = 19.2$ bristles.

7. (a) Total variance $= 3r^2/4$; additive variance $= r^2/2$; heritability $= 2/3$. (b) $S = 3rt^2/(4 + 3S)$; $R = s_r(8 + 5S)/(2(4 + 3S)^2)$; hence $h^2 = (8 + 5S)/(3(4 + 3S))$, which tends to $2/3$ as $s \to 0$.

8. Total variance $= 9$, and $h^2 = 0.3$, so $V_s = 2.7$. Total difference $D = 16$. From Box 6.6, if initial gene frequencies are 0.5, number of loci $= D^2/\delta V_s = 12$ loci.

9. $0.63ab : 0.37AB$. Note that this implies instability of the equilibrium shown in Table 6.5, because the frequencies have moved away from the equilibrium point. Usually, epistatic fitnesses give rise to alternative stable monomorphic states, with an unstable equilibrium between them. Stable polymorphic states, such as heterostyly and mimicry, are maintained by frequency-dependent fitnesses.

10. If there are $n$ loci, and the effect per allele substitution is $d$, then the initial variance is $n \times 2 \times 0.01 \times 0.99 d^2 = 0.0198 d^2$. The maximum variance is reached when half the loci have $p = q = 0.5$; by this time the remaining loci will contribute little. Hence the maximum variance $= (n/2) \times 2 \times 0.5^2 \times d^2 = 0.25 nd^2$. That is, the genetic variance is increased by a factor of 12.6. In practice, the increase would be less than this, because different loci would change at different rates, but it should still be detectable.

Chapter 7

1. (a) All $R$, or all $S$. (b) The unstable equilibrium, with $R$ and $S$ equally fit, occurs when $p(R) = 0.25$, $p(S) = 0.75$. If $R$ is initially commoner than 0.25, it will increase. When $p(A) = 0.4$, $p(a) = 0.6$, and hence, with random mating, $p(R) = 0.36$. Hence the population will evolve to $p(R) = 1$, $p(a) = 1$.

2. The ESS is $p(R) = 0.25$, $p(S) = 0.75$. Therefore $p^2(a) = 0.75$, or $p(a) = 0.866$.

3. There are two ESS’s: all $C$, or $M = (0.5A : 0.5B)$. Note that $E(M,M) = 3.5$, and $E(C,M) = 2$, so $C$ cannot invade $M$.

4. Let $FA$, $FL$, and $FS$ be the strategies ‘always escalate’, ‘escalate if larger, withdraw if smaller’, and ‘escalate if smaller, withdraw if larger’, respectively. In a contest, an individual has a 50 per cent chance of being larger. (a) $E(FA,FA) = R/2 - C/2$; $E(FL,FA) = [PR - (1 - P)C]$. Hence $FA$ is stable against invasion by $FL$ provided that $R - C > PR - (1 - P)C$, or $(1 - P)R > PC$. This is possible provided that $P \neq 1$, and the reward is large relative to the cost. (b) $E(FA,FS) = R/2$; $E(FA,FS) = [RP - (1 - P)C] + R/2$. Hence $FS$ is stable against invasion by $FA$ provided that $RP - (1 - P)C < 0$, or $RP < (1 - P)C$. Hence the paradoxical strategy, ‘escalate if smaller’, can be an ESS, provided that $P \neq 1$, and costs are large relative to rewards. But I doubt whether such a paradoxical strategy has often evolved. Whenever the payoffs are such as to make it possible, the ‘common-sense’ strategy ‘escalate if larger’ is also an ESS. (c) The ‘payoff’ for death means the change in fitness caused by death. Hence $C$ is the expected reproductive success of an animal that withdraws from the contest without fighting.