

their genes to their offspring. Therefore, a parent and its offspring have  $\frac{1}{2}$  of their genes in common. Now consider two siblings A and B. Every maternal gene that has been transmitted to A has a 50% chance of also being inherited to B. Hence, A and B share  $\frac{1}{2}$  of their maternal genes, on average, the same being valid for paternal genes. Therefore, the average fraction of genes that are identical by descent is  $\frac{1}{2}$  for two siblings. This is just the fraction of genes shared by a parent and its offspring. Similar considerations show that, on average, half siblings have  $\frac{1}{4}$  and first cousins have  $\frac{1}{8}$  of their genes in common.

Now let us investigate a more elaborate model, namely that of a quantitative character. Our aim is to develop some of the basic notions of quantitative genetics.

### 3.1 AVERAGE EFFECTS AND GENETIC VARIANCE

Quantitative characters are traits that exhibit continuous or almost continuous variation, and can be measured on a metric scale. Typical quantitative characters are weight, height, various morphological measurements, yield, or fitness. Usually such traits are influenced by a large number of loci, often with small effects. It is generally not useful to describe the population genetics of such characters by gene frequencies, because they are difficult or impossible to measure. Instead, a phenotypical approach and statistical methods are needed, and the description will be in terms of the distribution of the character. The most important quantities to describe a probability distribution are its mean value, measuring the location, and its variance, measuring the dispersion. Often, these quantities can be accurately estimated from data. In this section, we consider the simplest case, in which a character is determined by a single locus and environmental influences are disregarded. More realistic models will be investigated later in this book.

We consider a diploid population that is not necessarily in Hardy-Weinberg equilibrium, and denote by  $P_{ij}$  ( $= P_{ji}$ ) the frequency of the ordered genotype  $\mathcal{A}_i\mathcal{A}_j$  and by  $p_i$  the frequency of allele  $\mathcal{A}_i$  (2.4). Let  $G_{ij}$  be the *genotypic value* (the measurement) of  $\mathcal{A}_i\mathcal{A}_j$  individuals. Then the *mean genotypic value* (the population mean) is

$$\bar{G} = \sum_{i,j} G_{ij} P_{ij} . \quad (3.1)$$

The deviation of  $G_{ij}$  from the mean is denoted by

$$g_{ij} = G_{ij} - \bar{G} , \quad (3.2)$$

and is called the *average excess* of the genotype  $\mathcal{A}_i\mathcal{A}_j$ . The *genotypic variance*, or (*total*) *genetic variance*, is defined as

$$\sigma_G^2 = \sum_{i,j} g_{ij}^2 P_{ij} . \quad (3.3)$$

Although, genotypic values and the total genetic variance are fundamental and measurable quantities, they do not necessarily reflect the 'evolutionary potential' of a population. The reason is that parents pass on their genes to the next generation, but not their genotypic values.

**Example 1.** As an illustration, consider a population in Hardy–Weinberg proportions with two alleles  $\mathcal{A}_1$  and  $\mathcal{A}_2$ , at frequencies  $p$  and  $1 - p$ , respectively. Then the three possible genotypes have frequencies  $p^2$ ,  $2p(1 - p)$ , and  $(1 - p)^2$ . Let us assume complete dominance, i.e., the genotypic values of the genotypes are  $G_{11} \neq G_{12} = G_{22}$ , and consider two individuals, one with genotype  $\mathcal{A}_1\mathcal{A}_2$ , the other  $\mathcal{A}_2\mathcal{A}_2$ . Hence, they have identical genotypic values,  $G_{12} = G_{22}$ . Suppose that each of these individuals is mated randomly with a number of other individuals. What will be the expected genotypic value of each's offspring? With the help of Table 2.1, and remembering that the probabilities that a given individual mates randomly with an  $\mathcal{A}_1\mathcal{A}_1$ ,  $\mathcal{A}_1\mathcal{A}_2$ , or  $\mathcal{A}_2\mathcal{A}_2$  individual are  $p^2$ ,  $2p(1 - p)$ , or  $(1 - p)^2$ , respectively, we calculate the expected genotypic value,  $\bar{G}'_{12}$ , among the offspring of  $\mathcal{A}_1\mathcal{A}_2$  as follows:

$$\begin{aligned}\bar{G}'_{12} &= p^2(\tfrac{1}{2}G_{11} + \tfrac{1}{2}G_{12}) + 2p(1 - p)(\tfrac{1}{4}G_{11} + \tfrac{1}{2}G_{12} + \tfrac{1}{4}G_{22}) \\ &\quad + (1 - p)^2(\tfrac{1}{2}G_{12} + \tfrac{1}{2}G_{22}) \\ &= \tfrac{1}{2}pG_{11} + (1 - \tfrac{1}{2}p)G_{22} .\end{aligned}\tag{3.4a}$$

The expected genotypic value among offspring of the  $\mathcal{A}_2\mathcal{A}_2$  individual, however, is

$$\bar{G}'_{22} = G_{22} ,\tag{3.4b}$$

which differs from  $\bar{G}'_{12}$  unless  $p = 0$ . Thus, the expected genetic measurement of an offspring does not depend in a simple way on the genotypic value of its parent.  $\diamond$

Therefore, we wish to assign to each allele  $\mathcal{A}_i$  a value  $\gamma_i$  that measures the average contribution of  $\mathcal{A}_i$  to the character and represents the ‘part’ of the genotypic value that determines the “genetic potentiality of the individual” (Fisher 1958, p. 33). We adopt a procedure based on a least-squares approximation devised by Fisher (1918, 1930, 1941). Although this seems arbitrary at the moment, it leads to a natural formulation for the correlation between relatives, and paves the way for understanding the response of quantitative traits to selection. Let us write

$$g_{ij} = \gamma_i + \gamma_j + \vartheta_{ij} ,\tag{3.5}$$

where  $\gamma_i$  is called the *average effect* of  $\mathcal{A}_i$  on the character and  $\vartheta_{ij}$  is the *dominance deviation*. The alleles  $\mathcal{A}_i$  and  $\mathcal{A}_j$  are said to act *additively* (on the given scale) if  $\vartheta_{ij} = 0$ . Otherwise, one speaks of dominance of effects. The least-squares procedure suggested by Fisher requires the minimization of the *dominance variance*

$$\sigma_D^2 = \sum_{i,j} \vartheta_{ij}^2 P_{ij} = \sum_{i,j} (g_{ij} - \gamma_i - \gamma_j)^2 P_{ij}\tag{3.6}$$

with respect to the  $\gamma_i$ . The idea behind this is, of course, to approximate the genotypic value  $G_{ij}$  as closely as possible by a linear expression,  $\bar{G} + \gamma_i + \gamma_j$ , in the sense that the expected value of the squared deviations  $\vartheta_{ij}^2$  is minimized. Differentiation of  $\sigma_D^2$  with respect to  $\gamma_i$  leads to the conditions

$$\sum_j \vartheta_{ij} P_{ij} = \sum_j (g_{ij} - \gamma_i - \gamma_j) P_{ij} = 0 \quad \text{for every } i .\tag{3.7}$$

From the definition (3.2) of  $g_{ij}$ , (3.5), and (2.4), we obtain

$$0 = \sum_{i,j} g_{ij} P_{ij} = 2 \sum_i \gamma_i p_i + \sum_{i,j} \vartheta_{ij} P_{ij}.$$

Hence, (3.7) implies that the mean of the average effects is

$$\sum_i \gamma_i p_i = 0. \quad (3.8)$$

The average excess  $g_i$  of  $\mathcal{A}_i$  is defined by  $g_i p_i = \sum_j g_{ij} P_{ij}$  and, because of (3.5) and (3.7), satisfies the relation

$$g_i p_i = \gamma_i p_i + \sum_j \gamma_j P_{ij}. \quad (3.9)$$

Therefore, the average effects,  $\gamma_i$ , can be found as the solution of the system of linear equations given by (3.9). If all allele frequencies satisfy  $p_i > 0$ , then the average effects are uniquely determined subject to the condition (3.8); see Chapter II.3.6. It is important to note that average effect and average excess depend on the gene frequencies and thus on the current genetic composition of the population. Therefore, they may change from one generation to the next.

~~$$\text{information of the average effect of the average excess of } p_i \text{ is } \gamma_i p_i \text{ and } g_i p_i \text{ respectively.} \quad (3.9)$$~~

~~$$\text{information of the average effect of the average excess of } p_i \text{ is } \gamma_i p_i \text{ and } g_i p_i \text{ respectively.} \quad (3.10)$$~~

The part of the total genetic variance that can be accounted for by the average effects of the alleles is called the *additive genetic*, or *genic*, *variance*. In view of (3.8), it is defined as

$$\sigma_A^2 = \sum_{i,j} (\gamma_i + \gamma_j)^2 P_{ij}. \quad (3.11)$$

With Hardy-Weinberg proportions, simple algebra verifies that this is equivalent to the representation

$$\sigma_A^2 = 2 \sum_i \gamma_i^2 p_i. \quad (3.12)$$

Finally, we obtain the following fundamental decomposition of the (total) genetic variance

$$\begin{aligned} \sigma_G^2 &= \sum_{i,j} (\gamma_i + \gamma_j + \vartheta_{ij})^2 P_{ij} \\ &= \sum_{i,j} (\gamma_i + \gamma_j)^2 P_{ij} + \sum_{i,j} \vartheta_{ij}^2 P_{ij} + 2 \sum_{i,j} (\gamma_i + \gamma_j) \vartheta_{ij} P_{ij} \\ &= \sigma_A^2 + \sigma_D^2, \end{aligned} \quad (3.13)$$

where the last sum in the middle line is zero because of (3.7). The dominance variance  $\sigma_D^2$  is typically much less than the additive genetic variance  $\sigma_A^2$  and vanishes in the absence of dominance.

A trivial, but important, consequence of the Hardy-Weinberg Law is that under random mating the genetic variance and its components remain constant in the absence of forces such as selection, mutation, or random genetic drift.

**Example 2.** Let us illustrate the notions just introduced for the case of two alleles, assuming Hardy-Weinberg proportions. Denote the frequency of  $\mathcal{A}_1$  by  $p$  and the genotypic contributions of  $\mathcal{A}_1\mathcal{A}_1$ ,  $\mathcal{A}_1\mathcal{A}_2$ , and  $\mathcal{A}_2\mathcal{A}_2$  by  $G_{11}$ ,  $G_{12}$ , and  $G_{22}$ , respectively. Then the mean genotypic value becomes

$$\bar{G} = G_{22} + 2(G_{12} - G_{22})p + 2\vartheta p^2, \quad (3.14)$$

where  $\vartheta = \frac{1}{2}(G_{11} + G_{22}) - G_{12}$ . From (3.10), the average effects of  $\mathcal{A}_1$  and  $\mathcal{A}_2$  are calculated to be  $\gamma_1 = -(1-p)(G_{22} - G_{12} - 2\vartheta p)$  and  $\gamma_2 = p(G_{22} - G_{12} - 2\vartheta p)$ , respectively, and, from (3.5), the dominance deviations are  $\vartheta_{11} = 2\vartheta(1-p)^2$ ,  $\vartheta_{12} = -2\vartheta p(1-p)$ , and  $\vartheta_{22} = 2\vartheta p^2$ . Thus, the additive genetic variance works out to be

$$\begin{aligned} \sigma_A^2 &= 2p(1-p)\left[\frac{1}{2}(G_{22} - G_{11}) + \vartheta(1-2p)\right]^2 \\ &= 2p(1-p)(\gamma_1 - \gamma_2)^2, \end{aligned} \quad (3.15)$$

and the dominance variance reduces to

$$\sigma_D^2 = 4\vartheta^2 p^2(1-p)^2. \quad (3.16)$$

Under the assumption of random mating,  $\gamma_1 - \gamma_2$  is known as the (average) effect of allelic substitution. Returning to the example at the beginning of this subsection, but assuming arbitrary dominance relations, the expected genotypic value among offspring of an  $\mathcal{A}_1\mathcal{A}_2$  individual with genotypic value  $G_{12} = \bar{G} + \gamma_1 + \gamma_2 + \vartheta_{12}$  is calculated to be

$$\bar{G}'_{12} = \bar{G} + \frac{1}{2}(\gamma_1 + \gamma_2) \quad \text{and} \quad \bar{G}'_{22} = \bar{G} + \gamma_2. \quad (3.17a)$$

◇

Thus, the average offspring differs from the population mean by half the additive effect of its parent, irrespective of dominance effects. Therefore, the average effects may be considered to measure the genetic and evolutionary 'essential' properties of a genotype. We shall see later in this book that it is, indeed, the additive genetic variance that is the major determinant for the response of the mean genotypic value to selection. The sum of the average effects of an individual is called its *breeding value*. For instance,  $\gamma_1 + \gamma_2$  is the breeding value of  $\mathcal{A}_1\mathcal{A}_2$ . Since this is twice the expected deviation of its offspring mean genotypic value from the population mean, it is a readily measurable quantity. For an arbitrary number of alleles, (3.17a) can be generalized and becomes

$$\bar{G}'_{ij} = \sum_k (G_{ik} + G_{jk})p_k = \bar{G} + \frac{1}{2}(\gamma_i + \gamma_j). \quad (3.17b)$$

If the population is not in Hardy-Weinberg proportions, then the relations (3.17) become more complicated. For example, the second equation in (3.17a) has to be replaced by  $\bar{G}'_{22} = \bar{G} + \gamma_2 + \vartheta_{22}(P_{12}^2 - P_{11}P_{22})/P_{12}$ , as the reader is invited to prove.

### 3.2 CORRELATIONS BETWEEN RELATIVES

The appropriate measure for the expected resemblance between individuals with respect to a given quantitative character is the correlation coefficient. Suppose for some