

Quantitative trait loci and gene interaction: the quantitative genetics of metapopulations

CHARLES J. GOODNIGHT

Department of Biology, University of Vermont, Marsh Life Science Building, Burlington, VT 05405-0086, U.S.A.

Genetic population differentiation is typically viewed as differentiation of population means. However, several theories of evolution and speciation postulate that populations differentiate not only with respect to the population means, but also with respect to the effects of alleles within these populations. I develop herein a measure of population differentiation for the ‘local average effects’ of alleles, where local average effect is defined as the average effect of an allele in a deme measured as a deviation from the metapopulation mean. The differentiation for local average effects has two components, a component attributable to the population mean and a residual component that is attributable to changes in the local average effects independent of the population mean. The variance in local average effects attributable to the population mean is measured as the variance in the mean local average effect of all alleles. The variance in the residual local average effects is measured as the difference between the variance local average effects of individual alleles and the variance in the mean local average effects of all alleles. Differentiation for population means and differentiation for residual local average effects need not be related. I show that when there is only additive gene action, populations can differentiate for population means, but not for residual local average effects. However, if there is gene interaction then populations can also differentiate for local average effects of alleles. The consequence of this differentiation is that the local average effects of alleles change relative to each other such that an allele that is favoured by selection in one population may be removed by selection in other populations. I discuss the evolutionary consequences of differentiation for local average effects, and the interpretation of QTL data in light of this model.

Keywords: genetic drift, genetic variance components, metapopulation quantitative genetics, quantitative trait loci, speciation.

Introduction

Genetic population differentiation typically is viewed as differentiation of population means (Wright, 1952; Crow & Kimura, 1970). However, a variety of theories of evolution and speciation postulate that populations differentiate not only with respect to the population means, but also with respect to the effects of alleles within these populations. This is the basis of Wright’s (1931, 1977) shifting balance process, which postulates that epistasis and pleiotropy will lead to ‘multiple selective peaks’. A hallmark of this theory is that different alleles are favoured by selection in the ‘domains of attraction’ of different peaks. Similarly, theories of speciation resulting from periods of small population size, such as Mayr’s (1963) genetic revolutions at speciation model, Carson’s (1968) founder-flush theory of speciation, and Templeton’s (1980a,b) theory of

genetic transience, all rely on the idea that the selective value of genes will change as a population passes through a genetic bottleneck.

These authors all considered gene interaction to be ubiquitous, but held in tightly coadapted gene complexes as a result of stabilizing selection on gene interactions during development. Under this view, the selective value of a gene is a property not only of the gene, but also of the genetic background in which it is found. Because the selective values of genes are a function of all of the interacting loci, an evolutionary inertia develops that severely constrains the evolutionary pathways along which a population can evolve. Founder events and genetic drift are seen as mechanisms for disrupting the cohesion of the gene pool by changing gene frequencies randomly with respect to fitness, and potentially causing a population to move under the domain of influence of a new coadapted gene complex. When a population begins to evolve towards a new coadapted gene complex alleles that formerly conferred high fitness, and were at high

Correspondence. E-mail: charles.goodnight@uvm.edu

frequency, may be removed by selection. Other alleles that were formerly deleterious and rare may suddenly confer high fitness and be favoured by selection. It is for this reason Wright felt that 'creative' evolution was most likely to occur when there was an interaction between genetic drift, mutation, migration and selection (Wright, 1977).

It has been shown that additive genetic variance can increase as a result of genetic drift in systems, when forms of gene interaction including within-locus dominance (Robertson, 1952; Willis & Orr, 1993), additive-by-additive epistasis (Goodnight, 1983, 1988; Cockerham & Tachida, 1988), and a variety of forms of two-locus epistatic interactions (Cheverud & Routman, 1996) are present. It must be true that if additive genetic variance increases as a result of population bottlenecks in systems with gene interactions, then a statistical conversion of nonadditive effects into additive effects is occurring. Because these nonadditive effects arise from gene interactions it seems reasonable intuitively that an among-populations genetic differentiation of average effects will occur. Thus, these studies of the increase in additive genetic variance lend credence to the suggestion that population bottlenecks can change the average effects of alleles.

In an earlier model I suggested that the genetic differentiation of demes within a metapopulation could be measured by calculating the variance in the local breeding value of a single sire across a metapopulation (Goodnight, 1995). The advantage of the local breeding value approach used in that model is that it is a direct extension of standard quantitative genetics, and is amenable to quantitative genetic experimental techniques. With the development of quantitative trait loci (QTL) methodologies it is reasonable to develop similar measures for systems with molecularly localized genetic effects. Recently Cheverud & Routman (1995, 1996; Routman & Cheverud, 1997) have developed a modification of Cockerham's (1954) two-locus genetic model that can be readily applied to QTL studies. Here I extend the models of Cheverud and Routman to include the measures of genetic population differentiation developed in Goodnight (1995). In addition to providing measures of population differentiation for average effects in terms amenable to QTL studies this approach can be used to analyse all forms of one- and two-locus genetic interactions, as well as two-locus genotypic values derived from empirical QTL studies. Thus, although there are clear limitations to this modelling approach it provides an excellent complement to the approach originally developed by Cockerham (1954).

Model

Epistasis occurs when phenotypic differences among genotypes at one locus depend on which genotypes are

present at other loci. Cheverud & Routman (1996) define a quantitative measure of 'physiological epistasis' in a two-locus genetic model that is analogous to the arbitrarily assigned genotypic values in a single-locus system with dominance used by Falconer & Mackay (1996) following Hayman & Mather (1955). This definition of physiological epistasis does not depend on allele frequencies at the two loci in question and it is distinct from the contribution of epistasis to genetic variance components. Physiological epistasis contributes to all three genetic variance components, additive, dominance and epistatic (Cheverud & Routman, 1996).

I consider a phenotype determined by two loci, an *A* locus and a *B* locus, each with two alleles, labelled 1 and 2, with allele frequencies p_1 and p_2 at locus *A* and q_1 and q_2 at locus *B*. The genotypic values, defined as the mean phenotype for a particular two-locus genotype, are independent of allele frequencies at the loci in question and, for present purposes, also assumed to be independent of allele frequencies at other loci (no three-way or higher epistasis). Each locus, *A* and *B*, may have a direct effect on the phenotype in question and act epistatically to modify the effects of the other locus (Cheverud & Routman, 1996).

I define the local average effect of the *i*th allele at the *l*th locus in the *m*th deme, $\alpha_{(L)mli}$, to be the mean deviation from the metapopulation mean of an individual containing the allele, with the remaining genotype having come at random from the *m*th deme. Average effects were developed to describe the phenotypic effects of an allele averaged across population structure (Fisher, 1958). As such, average effects are of greatest use in unstructured random mating populations (Falconer, 1985). Local average effects are a natural extension of average effects that explicitly include the effects of population subdivision (Wade & Goodnight, 1998). Local average effects are analogous to the local breeding values defined by Goodnight (1995). In standard quantitative genetics, breeding value is the sum of the average effects. This simple relationship does not hold for local breeding values (defined as the mean value of the offspring of an individual mated in a deme measured as a deviation from the metapopulation mean; Goodnight, 1995) and local average effects. This is the result of two factors. First, the local average effects include the effects of the local deme, and a sum of local average effects will inappropriately add the deme mean multiple times. To correct this the deme mean must be subtracted from each of the local average effects and then added back to the local breeding value (which is also measured as a deviation from the metapopulation mean). Secondly, the process of inbreeding causes alleles at the two loci to become associated with each other. At higher inbreeding coefficients some interactions ('cis' interactions) are

reliably transmitted from parent to offspring. As a result, the sum of the local average effects does not equal the local breeding value of a sire unless the sire is completely outbred ($f=0$; see also Falconer, 1985).

To examine the differentiation of demes for genetic effects I use the among-demes variance in the local average effects. It is this variance that will describe the extent to which alleles are performing differently in the different genetic backgrounds. The variance in the single-allele local average effects is:

$$\text{Var } \alpha_{(L)ij} = \sum_m p_m \alpha_{(L)mij}^2 \left(\sum_m p_m \alpha_{(L)mij} \right)^2, \quad (1)$$

and the variance in mean local average effect of a locus is the variance in the mean local average effect of all alleles at that locus in a particular deme:

$$\text{Var } \alpha_{(L)i\bullet} = \sum_m p_m \left(\sum_j p_{\bullet j} \alpha_{(L)mij} \right)^2 \left(\sum_m p_m \sum_j p_{\bullet j} \alpha_{(L)mij} \right)^2. \quad (2)$$

This is the among-demes variance in the mean local average effect of the i th locus. This measure is an intraclass covariance in average effects that is similar to the among-demes covariance in the local breeding values used by Goodnight (1995). A final measure of interest is the intraclass correlation in local average effects, or the fraction of the variance in local average effects that is among populations:

$$\text{Corr } \alpha_{(L)i\bullet} = \frac{\text{Var } \alpha_{(L)i\bullet}}{\text{Var } \alpha_{(L)ij}}. \quad (3)$$

Because epistatic values contribute to the local average effects of alleles, they also contribute to additive genetic variance and the variances in the local average effects of alleles, and they may do so at any allele frequency. The contributions of epistasis to dominance and interaction variance are specified in Cheverud & Routman (1996).

Following Cheverud & Routman (1996) I measure the effects of population bottlenecks on the variance in the local average effect of an allele and in the variance in mean local average effect. I do this by calculating these variances among populations of constant finite size drawn from a parental population with intermediate allele frequencies at both loci. The frequency distribution of populations displaying various combinations of allele frequencies is obtained using a two-locus Markov

chain model (Crow & Kimura, 1970; Hartl & Clark, 1997). The elements of this transition matrix are given by:

$$T_{(i(k+1))(j(l+1))} = \binom{2N}{j} \binom{2N}{l} \left(\frac{i}{2N} \right)^j \left(1 - \left(\frac{i}{2N} \right) \right)^{2N-j} \times \left(\frac{k}{2N} \right)^l \left(1 - \left(\frac{k}{2N} \right) \right)^{2N-l}, \quad (4)$$

where $T_{(i(k+1))(j(l+1))}$ is the transition probability from state (i,k) (i alleles at the A locus and k alleles at the B locus) to state (j,l) , and N is the population size. As with the standard Markov chain model this can be used to project the distribution of two-locus gene frequencies forward from an initial starting condition.

I then calculate the local average effect of alleles in each population type with a specified set of two-locus allele frequencies, and use these local average effects weighted by the population-type frequencies, to calculate the variances.

Most QTL analyses are undertaken using highly differentiated parental populations that are mated to produce a hybrid population with a gene frequency of 0.5 at all segregating loci. In keeping with this, all of the examples given in this paper assume a starting gene frequency of 0.5 at both loci, even though the Markov chain model can be calculated based on any starting set of gene frequencies. In natural populations the initial gene frequencies will normally not be at 0.5, particularly for loci that influence phenotypes that are under selection.

Special forms of genetic variation

In systems such as that modelled here, with two interacting loci, the nine genotypic values can be divided into eight independent genetical effects. The one- and two-locus genetic effects (Cockerham, 1954; Hayman & Mather, 1955; Van der Veen, 1959; Cockerham & Zeng, 1996) for a gene frequency of 0.5 at both loci and at linkage equilibrium are shown in Table 1. These genetic effects correspond to the columns of the 'S' matrix of the design III model of Cockerham & Zeng (1996). Also listed in Table 1 are the genotypic values for the two examples discussed below. Using regression methods described below, any two-locus interaction at any gene frequency can be decomposed into a combination of these eight forms of interaction. Table 2 lists the residual local average effects and Table 3 lists the among-deme variance components for the pure forms of genetic variance and the examples discussed below when all alleles are at fixation.

Table 1 Genotypic values for the pure forms of two-locus genetic variation (Cockerham, 1954) and the two examples of mixed forms of variance discussed. With the exception of the teocinte/maize example values are given in whole numbers for clarity. These numbers can be divided by the genetic standard deviation to obtain effects with genotypic variances of 1

	A_1A_1 B_1B_1	A_1A_2 B_1B_1	A_2A_2 B_1B_1	A_1A_1 B_1B_2	A_1A_2 B_1B_2	A_2A_2 B_1B_2	A_1A_1 B_2B_2	A_1A_2 B_2B_2	A_2A_2 B_2B_2
Additive <i>A</i> locus	1	0	-1	1	0	-1	1	0	-1
Additive <i>B</i> locus	1	1	1	0	0	0	-1	-1	-1
Dominance <i>A</i> locus	0	1	0	0	1	0	0	1	0
Dominance <i>B</i> locus	0	0	0	1	1	1	0	0	0
Additive-by-additive	1	0	-1	0	0	0	-1	0	1
Additive-by-dominance	1	0	-1	-1	0	1	1	0	-1
Dominance-by-additive	1	-1	1	0	0	0	-1	1	-1
Dominance-by-dominance	-1	1	-1	1	-1	1	-1	1	-1
Directional dominance (<i>A</i> locus)	1	1	0	1	1	0	1	1	0
<i>PEDS</i> (%) (teocinte × maize)	0.0	0.0	0.8	0.3	0.2	1.3	1.1	0.4	7.3

Table 2 Residual average effects of alleles at fixation for the pure forms of genetic variance

		Add. <i>A</i>	Add. <i>B</i>	Dom. <i>A</i>	Dom. <i>B</i>	<i>A</i> × <i>A</i>	<i>A</i> × <i>D</i>	<i>D</i> × <i>A</i>	<i>D</i> × <i>D</i>	Dir. Dom.	<i>PEDS</i>
Freq(<i>A</i> ₁) = 1	$\alpha_{(L)A1} - \alpha_{(L)A^*}$	1/2	0	-1/2	0	1/2	1/2	1	-1	0	-0.15
Freq(<i>B</i> ₁) = 1	$\alpha_{(L)A2} - \alpha_{(L)A^*}$	-1/2	0	1/2	0	-1/2	-1/2	-1	1	0	0.15
	$\alpha_{(L)B1} - \alpha_{(L)B^*}$	0	1/2	0	-1/2	1/2	1	1/2	-1	0	0.00
	$\alpha_{(L)B2} - \alpha_{(L)B^*}$	0	-1/2	0	1/2	-1/2	-1	-1/2	1	0	0.00
Freq(<i>A</i> ₁) = 1	$\alpha_{(L)A1} - \alpha_{(L)A^*}$	1/2	0	-1/2	0	-1/2	1/2	-1	-1	0	-0.25
Freq(<i>B</i> ₁) = 0	$\alpha_{(L)A2} - \alpha_{(L)A^*}$	-1/2	0	1/2	0	1/2	-1/2	1	1	0	0.25
	$\alpha_{(L)B1} - \alpha_{(L)B^*}$	0	1/2	0	1/2	1/2	-1	1/2	1	0	-0.40
	$\alpha_{(L)B2} - \alpha_{(L)B^*}$	0	-1/2	0	-1/2	-1/2	1	-1/2	-1	0	0.40
Freq(<i>A</i> ₁) = 0	$\alpha_{(L)A1} - \alpha_{(L)A^*}$	1/2	0	1/2	0	1/2	1/2	-1	1	1/2	-0.40
Freq(<i>B</i> ₁) = 1	$\alpha_{(L)A2} - \alpha_{(L)A^*}$	-1/2	0	-1/2	0	-1/2	-1/2	1	-1	-1/2	0.40
	$\alpha_{(L)B1} - \alpha_{(L)B^*}$	0	1/2	0	-1/2	-1/2	-1	1/2	-1	0	0.35
	$\alpha_{(L)B2} - \alpha_{(L)B^*}$	0	-1/2	0	1/2	1/2	1	-1/2	1	0	-0.35
Freq(<i>A</i> ₁) = 0	$\alpha_{(L)A1} - \alpha_{(L)A^*}$	1/2	0	1/2	0	-1/2	1/2	1	1	1/2	-3.00
Freq(<i>B</i> ₁) = 0	$\alpha_{(L)A2} - \alpha_{(L)A^*}$	-1/2	0	-1/2	0	1/2	-1/2	-1	-1	-1/2	3.00
	$\alpha_{(L)B1} - \alpha_{(L)B^*}$	0	1/2	0	1/2	-1/2	1	1/2	1	0	-3.45
	$\alpha_{(L)B2} - \alpha_{(L)B^*}$	0	-1/2	0	-1/2	1/2	-1	-1/2	-1	0	3.45
Var($\alpha_{(L)A1} - \alpha_{(L)A^*}$)		0	0	1/4	0	1/4	0	1	1	1/16	1.41
Var($\alpha_{(L)B1} - \alpha_{(L)B^*}$)		0	0	0	1/4	1/4	1	0	1	0	2.28

Figure 1(a–e) are graphs of the among-deme variances as a function of inbreeding coefficient for the population mean, local average effect of a single allele, and mean local average effects for the different forms of genetic effects. The mean additive genetic variance is plotted in addition. These graphs were calculated using the effect matrices for the interaction in Table 1 divided by the genetic standard deviation, so that the total genetic variance in the ancestral population ($p_1 = q_1 = 0.5$) is 1. Population sizes of 2, 4, 8, 16, 32 and 64 individuals were used in the calculations. Population

sizes of eight and larger are connected by a line. Smaller populations (two and four) frequently deviate from these smooth curves, particularly in interactions involving dominance.

Additivity

With pure additive variance, genotypic differences at the two loci are independent, and the genotypic value is a simple sum of the number of *A*₂ or *B*₂ alleles (Table 1, Fig. 1a). The variance in local average effects, the

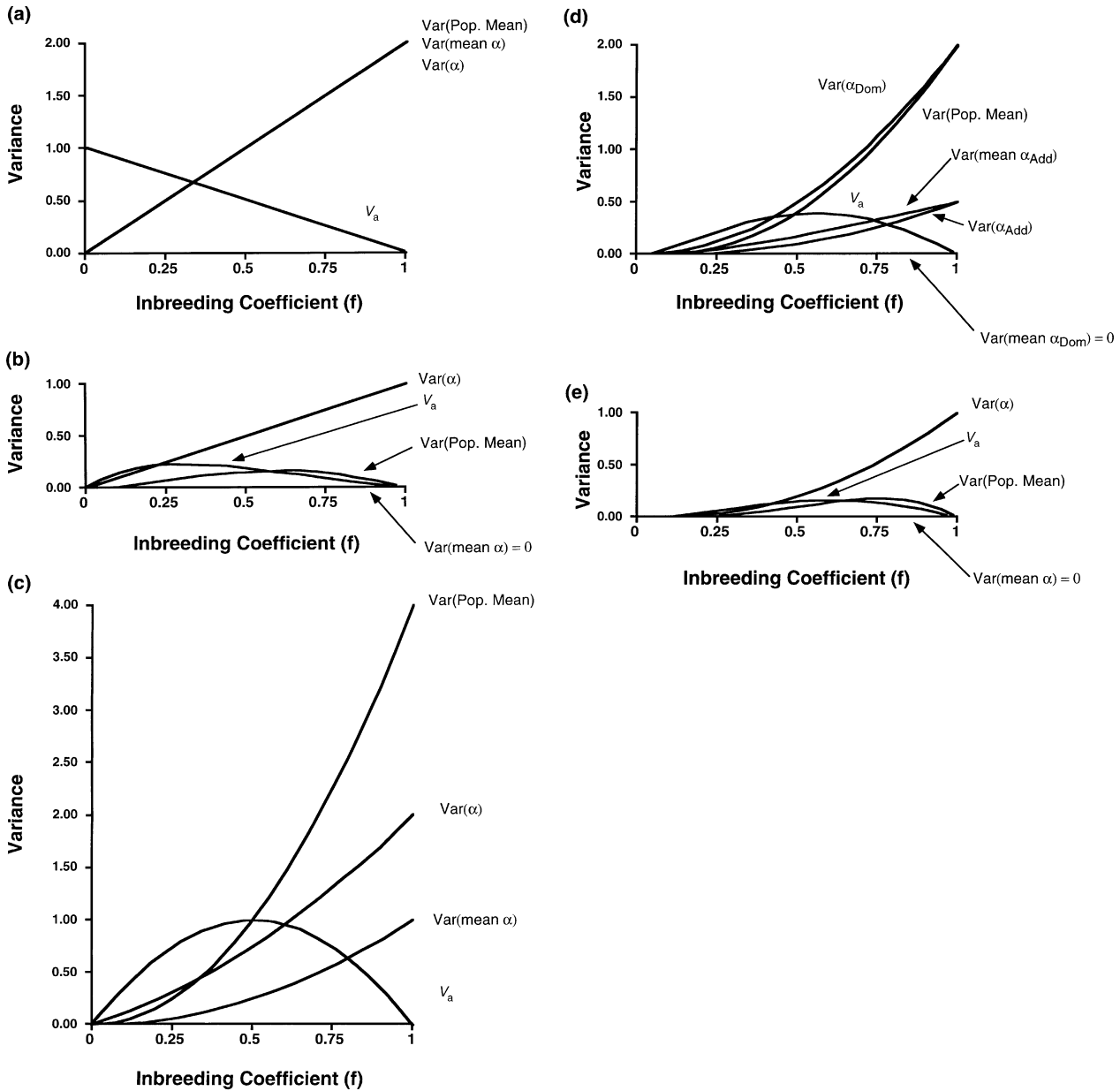


Fig. 1 The among-deme variance in population means ($\text{var}(\text{Pop. Mean})$), the variance in local average effect ($\text{Var}(\alpha)$), the variance in the mean local average effect ($\text{Var}(\text{mean}\alpha)$), and the additive genetic variance (V_a) as a function of inbreeding coefficient for the ‘pure’ forms of genetic effects. All examples are started with an initial gene frequency of 0.5 at both loci, and population sizes of 4, 8, 16, 32 and 64 individuals. In some cases smaller populations deviate slightly from this line. The difference ($\text{Var}(\alpha) - \text{Var}(\text{mean}\alpha)$) is the residual variance in average effects and measures the differentiation in local average effects. (a) Additive genetic variance (b) dominance genetic variance, (c) additive-by-additive epistasis, (d) additive-by-dominance epistasis, (the additive locus (Add) and the dominance locus (Dom) are plotted separately), (e) dominance-by-dominance epistasis.

variance in mean local average effects and the variance among populations are equal, and at fixation are equal to twice the additive variance (owing to that locus) in the ancestral population (Fig. 1a). As a result, the residual variance in local average effect is zero and the correlation in local average effects is 1.

Dominance

Pure dominance (Table 1, Fig. 1b) is what is classically called symmetrical overdominance. Later I consider directional dominance, which is a combination of dominance and additivity. With dominance the two

homozygotes have the same genotypic value, which is lower than the genotypic value of the heterozygote. As with the additive model the two loci are independent in their effects. Because the two homozygotes have equal genotypic values as the populations go to fixation, the variance among demes goes to 0. The variance in the average effect climbs to a value of 1.0. However, the mean average effect is constant for all population frequencies. Thus the variance in mean average effects is 0 and the correlation in average effects is also 0. This will be seen to be a general result: for all interactions involving dominance the variance in the mean local average effect, and, thus, the correlation in local average effects, is always 0.

Additive-by-additive epistasis

With additive-by-additive epistasis, genotypic differences at each locus are additive within modifying locus genotypes, but the additive effects differ in sign depending on which modifying locus genotype is present (Table 1, Fig. 1c). The additive-by-additive epistasis pattern described here corresponds to the special case examined in previous models (Goodnight, 1988, 1995; Whitlock *et al.*, 1993). The results of this model are numerically identical to those of a coancestry model (Goodnight, unpubl. data), thus Fig. 1(c) can be considered to be the results for a coancestry model as well.

Additive-by-dominance and dominance-by-additive epistasis

With additive-by-dominance epistasis (Table 1, Fig. 1d), phenotypic differences are additive at one locus (the *A* locus), but differ in direction depending on the genotype at the *B* locus. At the *B* locus, genotypic values display overdominance or underdominance depending on the genotype at the *A* locus. Dominance-by-additive epistasis is the transpose of additive-by-dominance epistasis, but it is otherwise identical.

For the additive locus (locus *A*) the correlation in local average effects starts out at 0, but the correlation goes to 1 as fixation occurs. This occurs because in the presence of either homozygous genotype at the *B* locus, the *A* locus is additive with the A_1 allele greater than the A_2 allele. At greater values of f there are relatively few *B* locus heterozygotes, and therefore relatively few populations showing a reversal in the ranking of the *A* alleles.

The dominance locus (*B*) behaves very differently from the additive locus. The variance in the local average effect of an allele is always greater than or equal to the variance among deme means, whereas the variance in the mean average effects is always 0. Thus, the correlation in local average effects is always 0. Of particular

interest is that the dominance locus is overdominant in the presence of the A_2 homozygote and underdominant in the presence of the A_1 homozygote (Table 1). Thus, for example, a homozygous B_1B_1 population is invisable by a B_2 allele in a population fixed for the A_2A_2 genotype, but not in a population fixed for the A_1A_1 genotype.

Dominance-by-dominance epistasis

With dominance-by-dominance epistasis (Table 1, Fig. 1e) all of the two-locus homozygotes have the same phenotype. However, heterozygote values vary so that genotypic values display underdominance or overdominance depending on the modifying locus genotype. As with the other dominance interaction, the variance in mean local average effect is 0. Throughout the inbreeding process the population means become only slightly differentiated, and at fixation all population means are the same ($\text{Var}(\mu)=0$). Note, however, that the local average effects of alleles are highly dependent on the gene frequencies so that the differentiation in average effects increases with increasing f (Table 3).

Mixtures of forms of genetic effects

The relative contributions of the different genetic effects of a pair of interacting loci (QTL) can be determined by performing a partial regression of the values in the observed genotypic values on the corresponding values in the pure forms of genetic variance. The partial regression as it is presented here is used strictly for partitioning the variances and not for significance testing. It has several details that distinguish it from regression as it is typically used. First, it is important that the type one or sequential sums of squares be used. Other types of sums of squares (e.g. type three sums of squares) use an iterative formula that will give incorrect results. Secondly, because sequential sums of squares are used it is important that the effects be entered in order of increasing degree of interaction. Thus, the pure forms of genetic variance should be entered as additive (one for each locus, order not important), dominance (one for each locus, order not important), additive-by-additive epistasis, additive-by-dominance and dominance-by-additive epistasis (order not important), and finally dominance-by-dominance epistasis. Thirdly, the pure forms of genetic variance (Table 2) should be divided by the standard deviation of that form of variance at the gene frequency being measured. This gives them a variance of one, which makes the interpretation of the regressions clearer. Finally, the regression should be weighted by the genotype frequencies. [A Macintosh JMP (SAS, 1994) spreadsheet for performing these regressions is available from the author.]

Table 3 Among-deme variance components at fixation

Source of genetic variance	Var (μ)	Var (α_{mij})	Var ($\alpha_{mi\cdot}$)	Corr ($\alpha_{mi\cdot}$)
Additive	2	1/2	1/2	1
Dominance	0	1	0	0
Additive-by-additive epistasis	4	2	1	1/2
Additive-by-dominance epistasis (additive locus)	2	1/2	1/2	1
Additive-by-dominance epistasis (dominant locus)	2	2	0	0
Dominance-by-dominance epistasis	0	1	0	0
Directional dominance	1 1/3	1/3	2/3	1/2
% of cupules with the pedicellate spikelet	8.495			
Locus <i>A</i> : UMC107			4.03	2.62
Locus <i>B</i> : BV302			4.60	2.32

This partial regression divides the empirical genotypic values into the pure additive, dominance and epistatic components with the sums-of-squares providing the variances attributable to the pure forms of genetic effects. Note that partitioning of the genotypic values into the pure forms of genetic variance will change as gene frequencies change. For example, what is expressed as epistasis in the F_2 population may be expressed as a mixture of additive and epistatic effects in the backcross populations. I illustrate this first with the theoretical example of directional dominance, and secondly with an empirical example involving two interacting QTL in a cross between teocinte and maize (Doebley *et al.*, 1995).

With mixtures of genetic effects that include both additive and dominance effects there will frequently be a covariance between the local average effect and the mean local average effect. For two alleles the covariance will be positive for the dominant allele and of equal magnitude but negative for the recessive allele. This covariance is removed when the mean variance in the local average effect of alleles is used (note that this is different from the variance in the mean local average effect of alleles).

Directional dominance

First I examine a theoretical interaction, directional dominance (Table 1, Fig. 2). In this example the *A* locus shows directional dominance with the ‘1’ alleles being dominant to the ‘2’ alleles and the *B* locus is neutral. Although often referred to as simply dominance, directional dominance is a combination of the additive and dominance effects described above. It is helpful to visualize directional dominance as a dominance interaction ‘tilted’ so that one of the homozygotes has a higher fitness than the other. The regression approach described above will divide this interaction into compo-

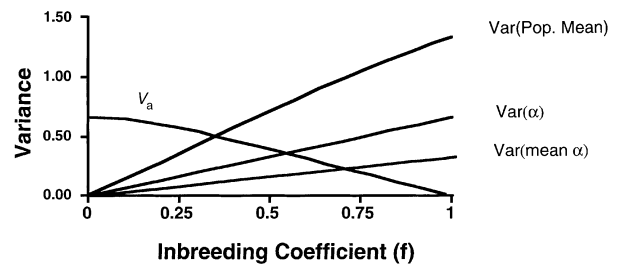


Fig. 2 The among-deme variance in population means (var(Pop. Mean)), the variance in local average effect (Var(α)), the variance in the mean local average effect (Var(mean α)), and the additive genetic variance (V_a) as a function of inbreeding coefficient for directional dominance. Initial gene frequency is 0.5, and population sizes of 2, 4, 8, 16, 32 and 64 individuals. Populations greater than four individuals are connected by a line.

nents attributable to dominance (the ‘overdominance’) and additivity (the ‘tilt’).

The correlation in local average effects remains 0.5 for all generations and population sizes and at fixation the only populations that have nonzero residual local average effects are those fixed for the recessive allele (Table 2). The local average effects of the dominant and recessive dominant alleles are both 0 in populations fixed for the dominant allele. In these populations a single recessive allele would inevitably be in a heterozygote with the dominant phenotype; thus, the neutrality of the recessive allele in these populations makes intuitive sense. On the other hand, in populations fixed for the recessive allele the local average effects of the two alleles are different and nonzero. Again, this is in accordance with expectation because heterozygotes would have the dominant phenotype unlike other members of the population that would have the homozygous recessive phenotype. Note that in this example it is the dominant phenotype that has high fitness, and

therefore could invade populations fixed for the recessive allele. However, the opposite is also possible with the recessive homozygote having the high fitness phenotype.

Performing a multiple regression on the eight pure forms of genetic variance indicates that a system with directional dominance is described fully by the additive and dominance variance. At a gene frequency of 1 for the recessive allele all of the genetic variance is additive, whereas at a gene frequency of 0 for the recessive allele all of the genetic variance is dominance variance. At an intermediate gene frequency ($P = 0.5$) $2/3$ of the variance is additive and $1/3$ is dominance variance (see Falconer & Mackay, 1996, p. 128).

Empirical example

Doebley *et al.* (1995) have identified QTL for several traits in a wide cross between teosinte (*Zea mays ssp. parviglumis*) and cultivated maize (*Zea mays ssp. mays*). The details of the mapping procedure are described in Doebley *et al.* (1995; see also Doebley & Stec, 1993). One pair of markers (UMC107 and BV302) located on different chromosomes was shown to interact epistatically for several traits. The trait I have chosen to focus on is 'PEDS', the percentage of cupules with the pedicellate (maize-like) spikelet (see Doebley *et al.*, 1995 for a more complete description of this trait).

The partitioning of genetic variance into the genetic variance components for PEDS is shown for three gene frequencies in Table 4. Also shown in this table is the percentage contribution of the genetic variance components, the total genetic variance, and the percentage of the total genetic variance that results from digenic epistasis. At all three gene frequencies the genetical effects are attributable mainly to additive and single-

locus dominance effects. However, in all cases digenic epistasis accounts for a substantial proportion (between 20.9% and 39.4%) of the total genetic variance. This is reflected in Fig. 3 which shows a substantial deviation from additivity both because the greatest level of additive genetic variance occurs at an intermediate inbreeding coefficient, and because there is a divergence between the variance in local average effects and the variance in mean local average effects. The genetic variance shifts dramatically as gene frequencies change (Table 4). For example, dominance-by-dominance epistasis is a negligible component of variance at a gene frequency of 0.25 for the teocinte alleles, but accounts for 15.4% of the genetic variance when these alleles are at a frequency of 0.75. Note, however, that this increased percentage contribution of dominance-by-dominance epistasis is caused by a decline in the total genetic variance rather than by an increase in the component of variance.

Discussion

The model presented by Goodnight (1995) and here extended to QTL is a metapopulation extension of quantitative genetics. This model shows that our standard methods of measuring dominance and epistasis within populations do not fully describe the role of gene interaction in evolutionary processes. Even when there is no apparent dominance or epistasis within populations, gene interactions may be causing differentiation of average effects among populations. Some of the most important consequences of QTL interactions are among populations and these are reflected in variance in the residual local average effects of alleles.

The implications of the residual variance in local average effects can be seen from Table 2. In additive

Table 4 Decomposition into genetic variance components for a pair of interacting loci (UMC107 and BV302) affecting the percentage of cupules lacking the pedicellate (maize-like) spikelet. Numbers in parentheses are the percentage of the total genetic variance due to the component. The total genetic variance, and the percentage due to digenic epistasis is also listed

Effect	Freq(<i>UMC107</i> _{teocinte}) = 0.25	0.5	0.75
	Freq(<i>BV302</i> _{teocinte}) = 0.25	0.5	0.75
Additive (UMC107)	3.7209 (38.6%)	0.6328 (21.6%)	0.0258 (8.9%)
Additive (BV302)	2.7719 (28.7%)	0.5513 (18.8%)	0.0697 (24.1%)
Dominance (UMC107)	0.8016 (8.3%)	0.4556 (15.5%)	0.0665 (23.0%)
Dominance (BV302)	0.3433 (3.5%)	0.1406 (4.8%)	0.0137 (4.7%)
A × A	1.4400 (14.9%)	0.4556 (15.5%)	0.0089 (3.1%)
A × D	0.3165 (3.3%)	0.3613 (12.3%)	0.0476 (16.4%)
D × A	0.2109 (2.2%)	0.1953 (6.7%)	0.0130 (4.5%)
D × D	0.0445 (0.5%)	0.1406 (4.8%)	0.0445 (15.4%)
Total genetic variance (% due to epistasis)	9.6501 (20.9%)	2.9331 (39.3%)	0.2897 (39.4%)

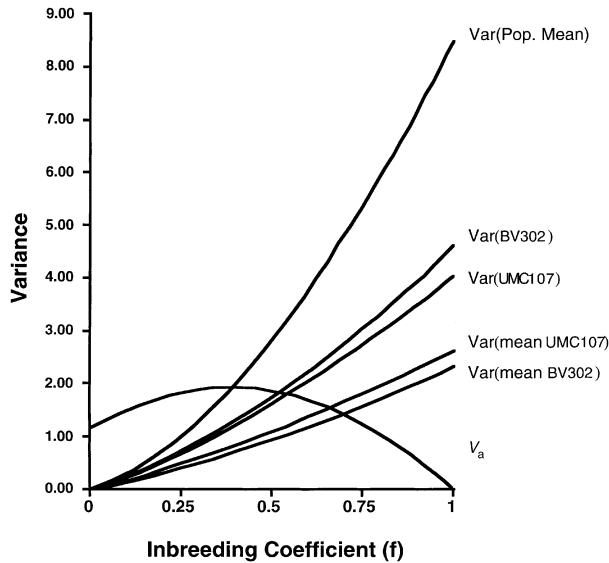


Fig. 3 The among-deme variance in population means ($\text{var}(\text{Pop. Mean})$), the variance in local average effect ($\text{Var}(\text{UMC107})$ and $\text{Var}(\text{BV302})$), the variance in the mean local average effect ($\text{Var}(\text{mean UMC107})$ and $\text{Var}(\text{mean BV302})$), and the additive genetic variance (V_a) as a function of inbreeding coefficient for the maize example. The initial gene frequency is 0.5, and population sizes of 8, 16, 32 and 64 individuals. Populations sizes smaller than eight individuals deviate slightly from this line.

systems with no dominance or epistasis this variance is 0. This means that the relative difference in the local average effects is a constant. For example, if, in one population, the difference in phenotype (say body weight) between individuals containing the A_1 vs. A_2 alleles is one gram, this difference will be maintained in all populations regardless of the frequency of alleles at the same or other loci. The local average effect of the individual alleles will change as a function of the mean of the population they are measured in. However, the difference between the two alleles will remain constant.

If there is dominance or epistasis the residual variance in local average effects will be nonzero. This means that the relative difference between local average effects of two alleles will not be constant. In this case, the difference in phenotype between individuals possessing the alternate alleles will vary among populations. In one population there may be a difference in phenotype of one gram, whereas a different population may have a larger difference (say two grams), and yet a third a smaller difference, or even a reversal of rank (say a difference of minus one gram). In the example of QTL alleles affecting the structure of the spikelets in the teocinte maize cross, in populations fixed for the B_1 allele the B_1 allele is either neutral or increases the

percentage of maize-like spikelets, whereas in populations fixed for the B_2 allele the B_1 allele decreases the percentage of maize-like spikelets (Table 2).

These results fit with the implications of Wright's shifting balance theory (Wright, 1931, 1977), as well as many of the models of founder event speciation (e.g. Mayr, 1963; Carson, 1968; Templeton, 1980a,b). These models all have in common the idea that there is extensive gene interaction, and that genetic drift or a population bottleneck leads to a change in the manner in which selection operates. This model indicates that when there is gene interaction, genetic drift or population bottlenecks can result in changes in the residual local average effects of alleles. These changes in the local average effects of alleles provide a potential mechanism for the effects postulated by these models.

Population differentiation is normally considered only in terms of its effect on population means (e.g. Hartl & Clark, 1997, pp. 498–501). Differentiation in the mean of a trait, as well as differentiation with respect to gene frequencies are both examples of differentiation of population means. This model (see also Goodnight, 1995) makes it clear that populations can also differentiate with respect to the local average effects of alleles. This form of population differentiation, differentiation for local average effects, need not be related to the differentiation of the population mean. At one extreme, systems with only additive effects can become differentiated for the population mean, but not for local average effects. At the other extreme, systems with pure dominance may not show any differentiation with respect to the population mean, but may become highly differentiated with respect to local average effects.

The idea that populations can differentiate for average effects is implicit in several models. It is an essential feature of Wright's shifting balance theory (1977), and it is implicit in several models of the genetics of speciation (Dobzhansky, 1936; Muller, 1939; Orr, 1995). Dempster (1963) observed that gene interaction can lead to shifts in the average effects of particular alleles, or even to reversals in sign as the genetic background alters. This model differs from these previous discussions in that it provides a quantitative measure of the differentiation for average effects that can be obtained experimentally.

For any pair of interacting QTL, regression can be used to divide the genetic variance into the different genetic effects. These quantitative genetic parameters are not constants. As gene frequencies change the genetic variance components will also change. In the maize example the amount of additive-by-additive epistatic variance varies from 14.9% to 3.1% as the frequency of teocinte genes changes from 0.25 to 0.75. Over this same range of frequencies the dominance-by-dominance epistatic variance increases from 0.5% to 15.4%. Similarly,

in the directional dominance example the partition of the total genetic variance between additive and dominance changes from being entirely dominance to entirely additive as the gene frequency moves from 0 to 1. This shifting between variance components is the basis for the changes in additive genetic variance reported for all of the pure forms of gene interaction. The pure forms of dominance and epistasis (Table 1) are only 'pure' at a gene frequency of 0.5 for both alleles.

It is instructive to consider some of the evolutionary implications of the different forms of epistasis. Two will be considered here. The first is the potential role of epistasis in speciation (Fig. 4). Consider a pair of QTL with the interaction specified for dominance-by-additive epistasis in Table 1. In a metapopulation fixed for the B_2 allele but segregating for the A locus, the A locus would be overdominant. Each subpopulation would stabilize at a gene frequency of 0.5 for the A_1 allele. If the subpopulations are finite, drift will cause some variation in this gene frequency; thus in any given subpopulation the actual gene frequency will be slightly different from 0.5. If a B_1 allele was introduced into a subpopulation, either through mutation or migration, the allele would be neutral only if the gene frequency of the A_1 allele was exactly 0.5. If it deviates from this in either direction the B_1 allele will be favoured by directional selection. Over time the B_1 allele is expected to go to fixation owing to a combination of genetic drift at the A locus and directional selection at the B locus. The result of this will be to shift the A locus from an overdominant locus to an underdominant locus. If the underdominance is large, or there are several pairs of dominance-by-additive loci, this could lead to reproductive isolation apparently attributable to the within-locus dominance interaction, when in fact it is the result of an epistatic interaction with a directionally selected locus. In many respects this is similar to the mechanism that Dobzhansky suggested was responsible for sterility in *Drosophila pseudoobscura* hybrids (Dobzhansky, 1936).

The second is the potential role of dominance-by-dominance epistasis in inbreeding depression (Fig. 5; see Crow & Kimura, 1970, pp. 77–81 and Schnell & Cockerham, 1992, for general discussions of the role of epistasis in inbreeding and heterosis). Note that in a population the interacting loci are neutral at a gene frequency of 0.5 for both loci, and this intermediate gene frequency will tend to be maintained by stabilizing selection. However, inbreeding will tend to cause the double homozygous genotypes to predominate. Because these genotypes all have relatively low fitnesses this will cause the general decline in fitness associated with inbreeding depression. This possibility is similar to models of inbreeding depression involving only dominance interactions such as overdominance or exposure

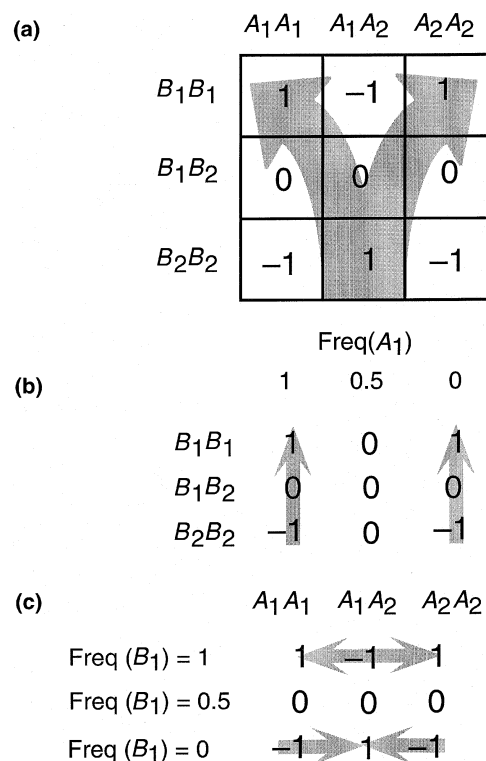


Fig. 4 The potential role of dominance-by-additive epistasis in speciation described in the text. (a) The genotypic values for dominance-by-additive epistasis. Grey arrows indicate the change from stabilizing selection to disruptive selection at the A locus that occurs as the frequency of the B_1 allele changes from zero to one. (b) Genotypic values for the three A locus genotypes when the frequency of the B_1 allele is 0, 0.5 and 1. When the B_1 allele is rare there is stabilizing selection at the A locus, whereas when the B_1 allele is common there is disruptive selection at the A locus. (c) Genotypic values for the three B locus genotypes when frequency of the A_1 allele is 0, 0.5 and 1. At a frequency at the A locus of 0.5 the B locus is neutral. However, if frequencies at the A locus drift from 0.5 then the B locus will be under directional selection favouring the B_1 allele. When the B_1 allele is rare, genetic drift at the A locus will interact with directional selection at the B locus eventually leading to the fixation of the B_1 allele, and either the A_1 or A_2 allele.

of deleterious alleles (Falconer & Mackay, 1996; Hartl & Clark, 1997), except that when viewed on a single-locus basis in the original outbred population the responsible loci appear to be neutral.

Fisher (1958) asserted that epistasis was effectively equivalent to environmental variance that could be ignored in the study of quantitative genetics. Fisher's assertion is a reasonable approximation for a single large randomly mating population. This model suggests that Fisher's assertion is not reasonable when populations are subdivided. The same forces that reduce the apparent contribution of genetic interactions to the

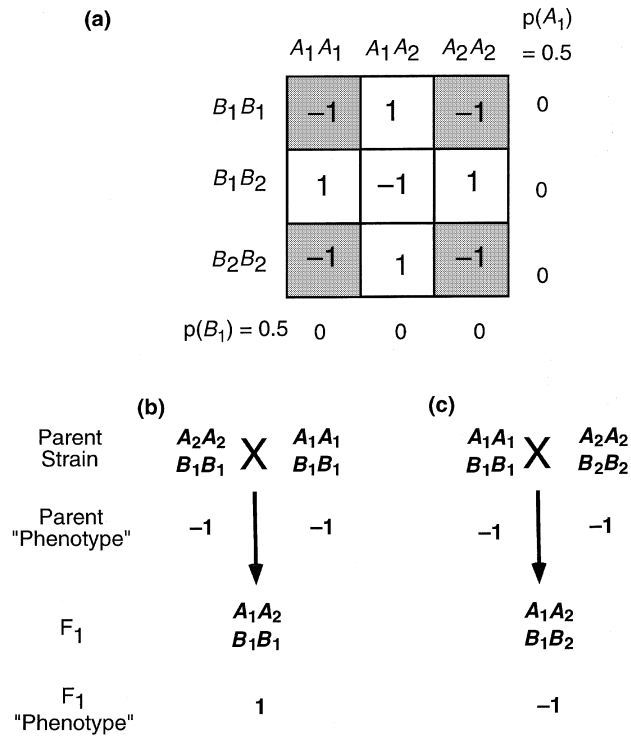


Fig. 5 The potential role of dominance-by-dominance epistasis in inbreeding depression described in the text. (a) The genotypic values for dominance-by-dominance epistasis. When both loci are at a frequency of 0.5 the two loci are neutral (one-locus genotypic values = 0). Inbreeding will increase the frequencies of the four double homozygous genotypes (shaded boxes) resulting in a decrease in fitness (inbreeding depression) even though there are no additive or dominance effects in the outbred population. (b) A cross between $A_1A_1B_1B_1$ and $A_2A_2B_1B_1$ results in heterosis (hybrid has higher fitness than either parent), whereas (c) a cross between $A_1A_1B_1B_1$ and $A_2A_2B_2B_2$ does not show heterosis.

variance within populations lead to populations differentiating for the local average effects of alleles. Among demes the effect of gene interactions is not to decrease the effectiveness of selection, as it does within demes; rather, it is to change the course of evolution. The metapopulation quantitative genetics outlined in this model are a straightforward extension of the basic quantitative genetic methods developed by Fisher. It is an extension that includes both standard quantitative genetic measures within populations and interdemic measures of population differentiation arising from gene interaction.

Acknowledgements

I thank J. Cheverud for suggesting this approach, helping me to understand the workings of his models, and for comments on early versions of the manuscript.

I also thank A. Templeton and M. Wade for helpful discussions and encouragement, and J. Doebley for making his data available to me. Much of the work on this project was performed while on sabbatical at Washington University in St. Louis in A. Templeton's lab. This work was supported by Vt EPSCoR, UVM grant BSCI95-3, and NSF DEB-9615260.

References

CARSON, H. 1968. The population flush and its genetic consequences. In: Lewontin, R. (ed.) *Population Biology and Evolution*, pp. 123-137. Syracuse University Press, Syracuse, NY.

CHEVERUD, J. M. AND ROUTMAN, E. J. 1995. Epistasis and its contribution to genetic variance components. *Genetics*, **139**, 1455-1461.

CHEVERUD, J. M. AND ROUTMAN, E. J. 1996. Epistasis as a source of increased additive genetic variance at population bottlenecks. *Evolution*, **50**, 1042-1051.

COCKERHAM, C. C. 1954. An extension of the concept of partitioning heredity variance for analysis of covariance among relatives when epistasis is present. *Genetics*, **39**, 859-882.

COCKERHAM, C. C. AND TACHIDA, H. 1988. Permanency of response to selection for quantitative characters in finite populations. *Proc. Natl. Acad. Sci., U.S.A.*, **85**, 1563-1565.

COCKERHAM, C. C. AND ZENG, Z. B. 1996. Design III with marker loci. *Genetics*, **143**, 1437-1456.

CROW, J. F. AND KIMURA, M. 1970. *An Introduction to Population Genetics Theory*. Burgess Publishing Co, Minneapolis, MN.

DEMPSTER, E. R. 1963. Concepts and definitions in relation to selection schemes. In: Hanson, W. D. and Robinson, H. F. (eds) *Statistical Genetics and Plant Breeding*, pp. 34-44. National Academy of Sciences/ National Research Council, Washington, DC.

DOBZHANSKY, T. 1936. Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*, **21**, 113-135.

DOEBLEY, J. AND STEC, A. 1993. Inheritance of the morphological differences between maize teosinte: comparison of results for two F₂ populations. *Genetics*, **134**, 559-570.

DOEBLEY, J., STEC, A. AND GUSTUS, C. 1995. Teosinte *branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics*, **141**, 333-346.

FALCONER, D. S. 1985. A note on Fisher's 'average effect' and 'average excess'. *Genet. Res.*, **46**, 337-347.

FALCONER, D. S. AND MACKAY, T. F. C. 1996. *Introduction to Quantitative Genetics*. Longman, Harlow, Essex.

FISHER, R. A. 1958. *The Genetical Theory of Natural Selection*, 2nd edn. Dover, New York.

GOODNIGHT, C. J. 1983. *Genetical and Environmental Factors Influencing Evolution by Group and Individual Selection*. Ph.D. Thesis, University of Chicago, IL.

GOODNIGHT, C. J. 1988. Epistasis and the effect of founder events on the additive genetic variance. *Evolution*, **42**, 441-454.

- GOODNIGHT, C. J. 1995. Epistasis and the increase in additive genetic variance: implications for phase I of Wright's shifting balance process. *Evolution*, **49**, 502–511.
- HAYMAN, B. I. AND MATHER, K. 1955. The description of genetic interactions in continuous variation. *Biometrics*, **11**, 69–82.
- HARTL, D. AND CLARK, A. G. 1997. *Principles of Population Genetics*, 3rd edn. Sinauer Associates, Sunderland, MA.
- MAYR, E. 1963. *Animal Species and Evolution*. The Belknap Press, Cambridge, MA.
- MULLER, H. J. 1939. Reversibility in evolution considered from the standpoint of genetics. *Biol. Rev. Camb. Phil. Soc.*, **14**, 261–280.
- ORR, H. A. 1995. The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, **139**, 1805–1813.
- ROBERTSON, A. 1952. The effect of inbreeding on the variation due to recessive genes. *Genetics*, **37**, 189–207.
- ROUTMAN, E. J. AND CHEVERUD, J. M. 1997. Gene effects on a quantitative trait: two-locus epistatic effects measured at microsatellite markers and at estimated QTL. *Evolution*, **51**, 1654–1662.
- SAS 1994. *JMP Statistics and Graphics Guide, Version 3*. SAS Institute, Inc., Cary, NC.
- SCHNELL, F. W. AND COCKERHAM, C. C. 1992. Multiplicative vs. arbitrary gene action in heterosis. *Genetics*, **131**, 461–469.
- TEMPLETON, A. R. 1980a. Models of speciation and inferences based on genetic distances. *Evolution*, **34**, 719–729.
- TEMPLETON, A. R. 1980b. The theory of speciation via the founder principle. *Genetics*, **94**, 1011–1038.
- VAN DER VEEN, H. J. 1959. Tests of non-allelic interaction and linkage for quantitative characters in generations derived from two diploid pure lines. *Genetica*, **30**, 201–232.
- WADE, M. J. AND GOODNIGHT, C. J. 1998. Genetics and adaptation in metapopulations: when nature does many small experiments. *Evolution*, **52**, 1537–1553.
- WHITLOCK, M. C., PHILLIPS, P. C. AND WADE, M. J. 1993. Gene interaction affects the additive genetic variance in subdivided populations with migration and extinction. *Evolution*, **47**, 1758–1769.
- WILLIS, J. H. AND ORR, A. H. 1993. Increased heritable variation following population bottlenecks: the role of dominance. *Evolution*, **47**, 949–956.
- WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics*, **16**, 93–159.
- WRIGHT, S. 1952. The genetics of quantitative variability. In: Reeve, E. C. R. and Waddington, C. H. (eds) *Quantitative Inheritance*, pp. 5–41. Her Majesty's Stationary Office, London.
- WRIGHT, S. 1977. *Evolution and the Genetics of Populations*, Vol. 3, *Experimental Results and Evolutionary Deductions*. University of Chicago Press, Chicago, IL.