The Contribution of Quantitative Trait Loci and Neutral Marker Loci to the Genetic Variances and Covariances Among Quantitative Traits in Random Mating Populations

Alfredo Ruiz* and Antonio Barbadilla†

**Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain, and †MCZ Laboratories, Harvard University, Cambridge, Massachusetts 02138*

Manuscript received July 15, 1994
Accepted for publication September 19, 1994

ABSTRACT

Using Cockerham's approach of orthogonal scales, we develop genetic models for the effect of an arbitrary number of multiallelic quantitative trait loci (QTLs) or neutral marker loci (NMLs) upon any number of quantitative traits. These models allow the unbiased estimation of the contributions of a set of marker loci to the additive and dominance variances and covariances among traits in a random mating population. The method has been applied to an analysis of allozyme and quantitative data from the European oyster. The contribution of a set of marker loci may either be real, when the markers are actually QTLs, or apparent, when they are NMLs that are in linkage disequilibrium with hidden QTLs. Our results show that the additive and dominance variances contributed by a set of NMLs are always *minimum* estimates of the corresponding variances contributed by the associated QTLs. In contrast, the apparent contribution of the NMLs to the additive and dominance covariances between two traits may be larger than, equal to or lower than the actual contributions of the QTLs. We also derive an expression for the expected variance explained by the correlation between a quantitative trait and multilocus heterozygosity. This correlation explains only a part of the genetic variance contributed by the markers, *i.e.*, in general, a combination of additive and dominance variances and, thus, provides only very limited information relative to the method supplied here.

THAT quantitative trait loci (QTLs) can be characterized using linked marker genes was first shown by SAX (1923) in the common bean and by THODAY (1961) in Drosophila. However, for many years, the low number of markers available was a limiting factor in these studies. The advent of techniques to detect molecular variation, beginning with protein electrophoresis and culminating in DNA sequencing, however, has reversed this situation: the huge amount of genetic variability revealed by these techniques in almost all species has allowed the construction of detailed genetic maps comprising hundreds of genetic markers evenly spaced throughout the genome and has made the mapping of QTLs feasible. Several maximum likelihood and regression methods have already been described for mapping QTLs using a variety of F_2, F_3, backcross and testcross generations (WELLER 1986, 1987; LANDER and BOTSTEIN 1989; KNAPP *et al.* 1990; HALEY and KNOTT 1992; MARTINEZ and CURNOV 1992; MORENO-GONZALEZ 1992), and their application has provided very promising results (TANKSLEY *et al.* 1982; EDWARDS *et al.* 1987; STUBER *et al.* 1987, 1992; PATERSON *et al.* 1988, 1991; ABLER *et al.* 1991; DEVICENTE and TANKSLEY 1993; NODARI *et al.* 1993).

Corresponding author: Alfredo Ruiz, Departament de Genètica i de Microbiologia, Edifici C, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain. E-mail: ibge5@cc.uab.es

The availability of an increasing number of genetic markers in many species also makes it possible to obtain information on the genetics of quantitative traits in outbred populations. One obvious difficulty when using genetic markers for this purpose is the possibility that they are not QTLs but neutral marker loci (NMLs) that are correlated with quantitative trait variation because they are in linkage disequilibrium with hidden QTLs. Using COCKERHAM's (1954) approach of orthogonal scales, we develop here multilocus genetic models for (1) the analysis of the contributions of a set of markers to the additive and dominance variances and covariances among quantitative traits and (2) the interpretation of these contributions when the markers are NMLs. Data amenable to such analysis are accumulating at an ever-faster rate, and a correct interpretation of the genetic information carried by markers is needed. The models allow the statistical estimation of the mentioned genetic contributions. As an example, we analyze data from the European oyster.

COCKERHAM (1954) showed how orthogonal scales could be used to partition the contribution of multiple loci to the genetic variance of a single trait into additive (A), dominance (D) and various (A × A, A × D, D

× D, . . .) interaction terms. COCKERHAM allowed for epistasis and inbreeding and assumed diallelic loci and linkage equilibrium. WEIR and COCKERHAM (1977) provided a complete partition of the genetic variance in a two-locus multiallelic system with inbreeding, linkage disequilibrium and arbitrary epistasis. Their results make it clear that the simultaneous consideration of all these factors in a multilocus system would yield mathematically untractable results. Here, COCKERHAM's approach is taken to build up a linear model for the effect of an arbitrary number of multiallelic QTLs upon any number of quantitative traits in a random mating population. The model assumes additive action across loci, *i.e.*, no epistasis, but allows for linkage disequilibrium, which may be important for reasons discussed below. We have also derived expressions for the apparent contributions made by an arbitrary number of multiallelic NMLs to the additive and dominance variances and covariances among any number of phenotypic traits. The results show that (1) a neat partition of the genetic variance into additive and dominance components is possible using NMLs, (2) the additive and dominance variances contributed by the NMLs are always *minimum* estimates of the corresponding variances contributed by the associated QTLs and (3) the contribution of the NMLs to the additive and dominance covariances may be larger than, equal to or lower than the actual contributions of the QTLs.

A number of studies in several species have found a correlation between growth rate, or other fitness-related traits, and multilocus heterozygosity estimated using allozyme markers (see for reviews, MITTON and GRANT 1984; ZOUROS and FOLTZ 1987). Although this correlation is evidence of the contribution (real or apparent) of the analyzed loci to the genetic variance of the traits, there is no clear interpretation of this relation. CHAKRABORTY and RYMAN (1983) noted that the correlation with multilocus heterozygosity may arise even with purely additive gene action. Using our multilocus model, we have derived an explicit expression for the genetic variance explained by the correlation with the multilocus heterozygosity. This has been done in two cases: first assuming that the studied loci are true QTLs and then assuming them to be NMLs. We show that the multilocus heterozygosity "captures" only a part of the total genetic variance contributed by the loci affecting the character, which is in general a combination of additive and dominance variances. Therefore, the correlation with multilocus heterozygosity provides very limited information relative to the method presented here.

RESULTS

Contribution of an arbitrary number of multiallelic QTLs to the genetic variances and covariances among phenotypic traits: The $k(k + 1)/2$ possible genotypes at a QTL with k alleles may be fully specified by means of $k - 1$ independent *allelic doses* (x_i) and $k(k - 1)/2$ *allelic dose products* (y_{ij}) (Table 1). These indicator variables are defined as

$$x_i = \xi_{im} + \xi_{if}, \quad (1)$$

$$y_{ij} = \xi_{im}\xi_{jf} + \xi_{jm}\xi_{if}, \quad (2)$$

where ξ_{im} and ξ_{if} (ξ_{jm} and ξ_{jf}) stand for the centered doses of allele i (j) in the male and female gametes, respectively, and are analogous to the first two orthogonal scales used by COCKERHAM (1954) to partition the genetic variance into components. Like COCKERHAM's scales, the allelic doses (x_i) are uncorrelated with the allelic dose products (y_{ij}) at the same or at a different locus (Table 2) and will allow us to decompose the genetic variances and covariances into additive and dominance components. Covariances among the allelic doses (or among the allelic dose products) within each locus are functions of the allelic frequencies, whereas similar covariances among different loci depend on gametic disequilibria (Table 2). One-locus disequilibria, digenic nongametic disequilibria and trigenic and quadrigenic disequilibria do not appear in these formulas because they are zero in a random mating population (WEIR and COCKERHAM 1989; WEIR 1990). Three-locus and higher order disequilibria (WEIR 1990; ROBINSON *et al.* 1991a) may be present, but, under the assumption of no epistasis, they do not contribute to the genetic variances and covariances.

When l multiallelic QTLs with a grand total of a alleles are considered, to specify the g possible one-locus genotypes, we require n independent allelic doses ($n = a - l$; one allele is left out at each locus) and m allelic dose products ($m = g - a$; the number of one-locus heterozygotes). Thus, we can define a ($n \times 1$) vector of independent allelic doses, **x**. We also have a ($n \times$

TABLE 1

Allelic doses (x_i) and allelic dose products (y_{ij}) for the different genotypes at a multiallelic locus in a random mating population

Genotype	Frequency	x_i	y_{ij}
$A_i\,A_i$	p_i^2	$2(1 - p_i)$	$2p_ip_j - 2p_j$
$A_i\,A_j$	$2p_ip_j$	$1 - 2p_i$	$2p_ip_j + (1 - p_i - p_j)$
$A_j\,A_j$	p_j^2	$-2p_i$	$2p_ip_j - 2p_i$
$A_i\,A_-$	$2p_i(1 - p_i - p_j)$	$1 - 2p_i$	$2p_ip_j - p_j$
$A_j\,A_-$	$2p_j(1 - p_i - p_j)$	$-2p_i$	$2p_ip_j - p_i$
$A_-\,A_-$	$(1 - p_i - p_j)^2$	$-2p_i$	$2p_ip_j$

A_i and A_j ($i \neq j$) are any two alleles at this locus (frequencies p_i and p_j, respectively); A_- stands for the remaining alleles (overall frequency = $1 - p_i - p_j$).

TABLE 2

Variances and covariances among the allelic doses and the allelic dose products of two multiallelic loci

Variances and covariances	x_i	y_{ij}
Same locus		
x_i	$2p_i(1 - p_i)$	0
x_j	$-2p_ip_j$	0
y_{ij}	0	$4p_i^2p_j^2 + 2p_ip_j(1 - p_i - p_j)$
y_{ir}	0	$-2p_ip_jp_r(1 - 2p_i)$
y_{rs}	0	$4p_ip_jp_rp_s$
Another locus		
x_i	$2D_{ii}$	0
y_{ij}	0	$2D_{ii}D_{jj} + 2D_{ij}D_{ji}$

D is the gametic disequilibrium parameter (LEWONTIN and KOJIMA 1960; WEIR 1990).

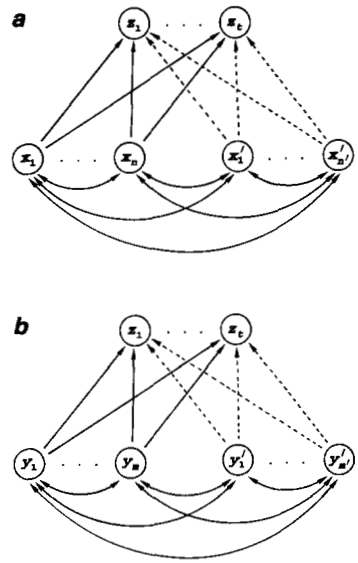

FIGURE 1.—Path diagrams showing the actual effects on t quantitative traits of a set of QTLs (continuous single-headed arrows) and the apparent effects of a set of NMLs correlated with them (discontinuous single-headed arrows). (a) Effects of the allelic doses (x and x' for the QTLs and the NMLs, respectively). (b) Effects of the allelic dose products (y and y' for the QTLs and the NMLs, respectively). The two effects are shown in separate graphs because the allelic doses and the allelic dose products are uncorrelated. Environmental effects have been omitted.

n) matrix of variances-covariances among the n allelic doses, $\mathbf{X}$, and a ($n \times n$) matrix of correlations among the n allelic doses, $\mathbf{R_x}$. The latter two matrices are related by the expression

$$\mathbf{X} = \mathbf{D_xR_xD_x}, \tag{3}$$

where $\mathbf{D_x}$ is a ($n \times n$) diagonal matrix with standard deviations of the allelic doses on the principal diagonal and zeros elsewhere. Likewise, a ($m \times 1$) vector, $\mathbf{y}$, a ($m \times m$) matrix of variances-covariances, $\mathbf{Y}$, and a ($m \times m$) matrix of correlations, $\mathbf{R_y}$, may be defined for the m allelic dose products. Of course, $\mathbf{Y}$ and $\mathbf{R_y}$ are related in the same way as $\mathbf{X}$ and $\mathbf{R_x}$ in Equation 3.

If the l QTLs affect t different measurable phenotypic traits, assuming no genotype–environment interaction and no epistasis, we can write the following model:

$$\mathbf{z} = \mathbf{a} + \mathbf{d} + \mathbf{e} = \boldsymbol{\alpha}^T\mathbf{x} + \boldsymbol{\delta}^T\mathbf{y} + \mathbf{e}, \tag{4}$$

where $\mathbf{z}$ and $\mathbf{a}$ are ($t \times 1$) vectors of phenotypic and additive values (given as deviations from the population mean); $\mathbf{d}$ and $\mathbf{e}$ are ($t \times 1$) vectors of dominance and environmental deviations, respectively; $\boldsymbol{\alpha}$ is a ($n \times t$) matrix of the average effects of the n alleles on the t phenotypic traits, expressed as deviations from the average effect of the allele left out in each locus, *i.e.*, the average effect of a gene substitution in case of a diallelic locus; $\boldsymbol{\delta}$ is a ($m \times t$) matrix of dominance parameters (deviations of the heterozygotes from the midpoint of the corresponding two homozygotes); and T stands for transpose.

This model can be represented by the path diagram shown in Figure 1 (for a description of path analysis, see WRIGHT 1921, 1934, 1968; LI 1975; for a multivariate generalization, see VOGLER and FULKER 1988). From Figure 1a, we derive the following expression for the matrix of additive variances-covariances:

$$\mathbf{A} = \mathbf{D_zp_{xz}^TR_xp_{xz}D_z}, \tag{5}$$

where $\mathbf{p_{xz}}$ is a ($n \times t$) matrix of path coefficients from the n allelic doses to the t phenotypic traits, equal to

$$\mathbf{p_{xz}} = \mathbf{D_x}\boldsymbol{\alpha}\mathbf{D_z^{-1}}, \tag{6}$$

and $\mathbf{D_z}$ is a ($t \times t$) diagonal matrix with phenotypic standard deviations of the t characters on the principal diagonal and zeros elsewhere. Likewise, from Figure 1b, we derive the following expression for the matrix of dominance variances-covariances:

$$\mathbf{D} = \mathbf{D_zp_{yz}^TR_yp_{yz}D_z}, \tag{7}$$

where $\mathbf{p_{yz}}$ is a ($m \times t$) matrix of path coefficients from the m allelic dose products to the t phenotypic traits

$$\mathbf{p_{yz}} = \mathbf{D_y}\boldsymbol{\delta}\mathbf{D_z^{-1}}. \tag{8}$$

Expressions 5 and 7 may be considered the multivariate generalization of formulas given by AVERY and HILL (1977, 1979), EWENS (1979) and WEIR *et al.* (1980).

Apparent contribution of an arbitrary number of multiallelic NMLs to the genetic variances and covariances among phenotypic traits: Let us consider now an arbitrary number of NMLs that have no effect on the quantitative traits *per se* but show an apparent effect due to their association with the QTLs. The characterization of the NMLs is fully analogous to that of the QTLs.

Thus, the same symbols are used for their allelic doses and allelic dose products but a prime is added to distinguish them from those of the QTLs. In addition, we define $\mathbf{C}_{xx'}$ and $\mathbf{R}_{xx'}$ as the $(n \times n')$ matrices of covariances and correlations, respectively, between the allelic doses of the QTLs and those of the NMLs. These two matrices are related by the expression:

$$\mathbf{C}_{xx'} = \mathbf{D}_x\mathbf{R}_{xx'}\mathbf{D}_{x'}, \quad (9)$$

where the meaning of $\mathbf{D}_x$ is given above and $\mathbf{D}_{x'}$ is the analogous diagonal matrix for the allelic doses of the NMLs. Similarly, we define $\mathbf{C}_{yy'}$ and $\mathbf{R}_{yy'}$ as the $(m \times m')$ matrices of covariances and correlations, respectively, between the allelic dose products of the QTLs and those of the NMLs. These two matrices are related in the same way as $\mathbf{C}_{xx'}$ and $\mathbf{R}_{xx'}$ in Equation 9.

We first derive an expression for the apparent contribution of the NMLs to the additive variances-covariances of the t phenotypic traits. The $(n' \times t)$ matrix of path coefficients from the allelic doses of the NMLs to the t phenotypic trait values (Figure 1a) is

$$\mathbf{p}_{x'z} = \mathbf{R}_{x'}^{-1}\,\mathbf{R}_{x'z}, \quad (10)$$

where $\mathbf{R}_{x'z}$ is the $(n' \times t)$ matrix of correlation coefficients between the allelic doses of the NMLs and the t phenotypic trait values:

$$\mathbf{R}_{x'z} = \mathbf{R}_{xx'}^{T}\mathbf{p}_{xz}. \quad (11)$$

After substitution of Equation 11 into Equation 10 we get

$$\mathbf{p}_{x'z} = \mathbf{R}_{x'}^{-1}\mathbf{R}_{xx'}^{T}\mathbf{p}_{xz}. \quad (12)$$

These path coefficients are apparent, *i.e.*, they are the standardized partial regression coefficients that would be observed were a multiple regression of the trait values on the allelic doses of the NMLs to be carried out. Now, the $(t \times t)$ matrix of variances-covariances "explained" by the n' allelic doses of the NMLs can be derived as

$$\mathbf{A}' = \mathbf{D}_z\mathbf{p}_{x'z}^{T}\mathbf{R}_{x'}\mathbf{p}_{x'z}\mathbf{D}_z. \quad (13)$$

After substitution of Equation 12 into Equation 13 we obtain

$$\mathbf{A}' = \mathbf{D}_z\mathbf{p}_{xz}^{T}\mathbf{R}_{xx'}\mathbf{R}_{x'}^{-1}\mathbf{R}_{xx'}^{T}\mathbf{p}_{xz}\mathbf{D}_z. \quad (14)$$

We are interested in the comparison of the true genetic variances-covariances contributed by the QTLs (expression 5), with the apparent contribution of the NMLs (expression 14). Subtracting Equation 14 from Equation 5, we get

$$\mathbf{A} - \mathbf{A}' = \mathbf{D}_z\mathbf{p}_{xz}^{T}\,(\mathbf{R}_x - \mathbf{R}_{xx'}\mathbf{R}_{x'}^{-1}\mathbf{R}_{xx'}^{T})\,\mathbf{p}_{xz}\mathbf{D}_z. \quad (15)$$

The meaning of this expression can be better understood if we take into account that

$$\mathbf{D}_x\mathbf{R}_{xx'}\mathbf{R}_{x'}^{-1}\mathbf{R}_{xx'}^{T}\mathbf{D}_x \quad (16)$$

is the $(n \times n)$ matrix of variances-covariances among the allelic doses of the QTLs jointly "explained" by the allelic doses of the NMLs. Thus the term

$$(\mathbf{R}_x - \mathbf{R}_{xx'}\mathbf{R}_{x'}^{-1}\mathbf{R}_{xx'}^{T}) \quad (17)$$

in expression 15 is a symmetric matrix whose diagonal elements represent the proportions of the variances of the allelic doses of the QTLs "unexplained" by the NMLs. These terms are thus all larger than or equal to zero. On the other hand, the terms off-diagonal represent the covariances between the allelic doses of the QTLs left "unexplained" by the NMLs divided by the product of the respective standard deviations and may thus be positive or negative. In conclusion, the apparent contribution of the NMLs to the additive variances of the phenotypic traits is always less than or equal to the true additive variances contributed by the QTLs. In contrast, the apparent contribution of the NMLs to the additive covariances between phenotypic traits may be higher or lower than the true contribution of the QTLs.

A similar argument can be used to derive an expression for the apparent contribution of the NMLs to the dominance variances-covariances of the t phenotypic traits. The $(t \times t)$ matrix of variances-covariances "explained" by the allelic dose products of the NMLs is (Figure 1b)

$$\mathbf{D}' = \mathbf{D}_z\mathbf{p}_{y'z}^{T}\mathbf{R}_{y'}\mathbf{p}_{y'z}\mathbf{D}_z, \quad (18)$$

where $\mathbf{p}_{y'z}$ is the $(m' \times t)$ matrix of path coefficients from the m' allelic dose products of the NMLs to the t phenotypic traits:

$$\mathbf{p}_{y'z} = \mathbf{R}_{y'}^{-1}\mathbf{R}_{yy'}^{T}\mathbf{p}_{yz}. \quad (19)$$

After substitution of Equation 19 into Equation 18, we get

$$\mathbf{D}' = \mathbf{D}_z\mathbf{p}_{yz}^{T}\mathbf{R}_{yy'}\mathbf{R}_{y'}^{-1}\mathbf{R}_{yy'}^{T}\mathbf{p}_{yz}\mathbf{D}_z. \quad (20)$$

The comparison between the apparent contribution of the NMLs to the dominance variances-covariances of the phenotypic traits and the true contribution of the QTLs is formally analogous to that carried out for the additive variances-covariances. The conclusions reached there hold true here as well. Note, however, that because the covariances between the allelic dose products of different loci are of the order of D^2 whereas those between the allelic doses are of the order of D (Table 2), we expect the apparent contributions to the dominance variances-covariances to be lower than those to the additive variances-covariances.

Genetic variance of a quantitative trait explained by the multilocus heterozygosity: Let us first assume that the studied loci are QTLs. The multilocus heterozygosity (the number of loci in the heterozygous state in each

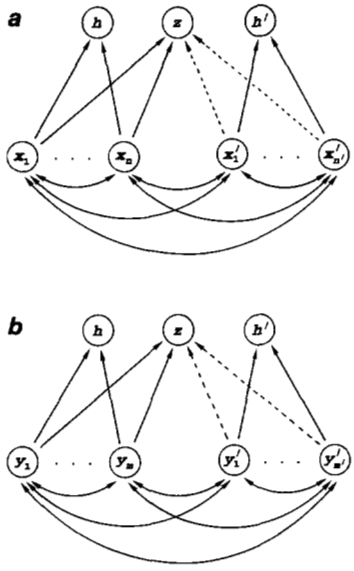

FIGURE 2.—Path diagram showing the relation between a quantitative trait (z) and the multilocus heterozygosity of a set of QTLs (h) or that of a set of NMLs (h'). All other symbols as Figure 1.

individual) may be treated as a quantitative trait without environmental variance, for which $\boldsymbol{\alpha}$ is a $(n \times 1)$ vector, $\mathbf{p}$, with element i equal to frequency of the excluded allele at the locus minus the frequency of the i allele, and $\boldsymbol{\delta}$ is a $(m \times 1)$ vector, $\mathbf{1}$, whose elements are all unity. The correlation between a quantitative trait (z) and the multilocus heterozygosity (h), assuming no environmental contribution, is then (Figure 2)

$$r_{zh} = \mathbf{p}_{xz}^{T}\mathbf{R}_{x}\mathbf{p}_{xh} + \mathbf{p}_{yz}^{T}\mathbf{R}_{y}\mathbf{p}_{yh}, \tag{21}$$

where $\mathbf{p}_{xh}$ and $\mathbf{p}_{yh}$ are, respectively, the $(n \times 1)$ and $(m \times 1)$ vectors of path coefficients from the allelic doses and the allelic dose products of the QTLs to the multilocus heterozygosity. The variance of the trait explained by h is thus

$$r_{zh}^{2}\sigma_{z}^{2} = \frac{\sigma_{z}^{2}}{\sigma_{h}^{2}}\,[\mathbf{p}_{xz}^{T}\mathbf{R}_{x}\mathbf{D}_{x}\mathbf{p} + \mathbf{p}_{yz}^{T}\mathbf{R}_{y}\mathbf{D}_{y}\mathbf{1}] \times [\mathbf{p}_{xz}^{T}\mathbf{R}_{x}\mathbf{D}_{x}\mathbf{p} + \mathbf{p}_{yz}^{T}\mathbf{R}_{y}\mathbf{D}_{y}\mathbf{1}]^{T}. \tag{22}$$

This expression shows that the variance explained by the multilocus heterozygosity is always less than, or equal to, the contribution of the QTLs to the genetic variance of the trait. Furthermore, in general, it does not represent either the additive or the dominance component but some combination of both. Only in the particular case when there is equality of allele frequencies within each locus and all the dominance parameters are identical, does the variance explained by the multilocus heterozygosity equal the dominance variance.

We now consider the case in which the multilocus heterozygosity (h') refers to a set of NMLs. The correlation between the trait value and h' may be derived (Figure 2) as follows:

$$r_{zh'} = \mathbf{p}_{xz}^{T}\mathbf{R}_{xx'}\mathbf{p}_{x'h'} + \mathbf{p}_{yz}^{T}\mathbf{R}_{yy'}\mathbf{p}_{y'h'}, \tag{23}$$

where the symbols follow the previous notation (the prime refers to the NMLs). The variance of the trait explained by h' is then given by the expression

$$r_{zh'}^{2}\sigma_{z}^{2} = \frac{\sigma_{z}^{2}}{\sigma_{h'}^{2}}\,[\mathbf{p}_{xz}^{T}\mathbf{R}_{xx'}\mathbf{D}_{x'}\mathbf{p}' + \mathbf{p}_{yz}^{T}\mathbf{R}_{yy'}\mathbf{D}_{y'}\mathbf{1}'] \times [\mathbf{p}_{xz}^{T}\mathbf{R}_{xx'}\mathbf{D}_{x'}\mathbf{p}' + \mathbf{p}_{yz}^{T}\mathbf{R}_{yy'}\mathbf{D}_{y'}\mathbf{1}']^{T}. \tag{24}$$

The variance of the trait explained by h' may be larger than, equal to or lower than that explained by h but is always less than, or equal to, the apparent contribution of the NMLs to the genetic variance of the trait, and only under very restrictive conditions does it amount to the dominance component.

Estimation of the contribution to the genetic variances and covariances: Consider a random sample of N individuals from a panmictic population in which each individual has been genotyped for l multiallelic marker loci and measured for t quantitative traits. According to the model given by expression 4, the observed data can be written in the form

$$\mathbf{Z}_{N} = \mathbf{X}_{N}\boldsymbol{\alpha} + \mathbf{Y}_{N}\boldsymbol{\delta} + \mathbf{E}_{N}, \tag{25}$$

where $\mathbf{Z}_{N}$ is the $(N \times t)$ data matrix whose row i comprises the observed values of the t measured quantitative traits on individual i; $\mathbf{X}_{N}$ and $\mathbf{Y}_{N}$ are, respectively, $(N \times n)$ and $(N \times m)$ matrices of individual allelic doses and allelic dose products (Table 1); $\mathbf{E}_{N}$ is a $(N \times t)$ matrix of random variables whose rows are independent observations from a multivariate normal distribution with mean zero and matrix of variances-covariances, $\mathbf{E}$. All other symbols are as above. A multivariate regression combines t multiple regressions into a single analysis, allowing the complete estimation of the unknown parameters, $\boldsymbol{\alpha}$, $\boldsymbol{\delta}$ and $\mathbf{E}$, and tests for the joint distribution of the model and for correlations (or covariances) among multiple regression lines. The maximum likelihood estimators of the unknown parameters are

$$\hat{\boldsymbol{\alpha}} = (\mathbf{X}_{N}^{T}\mathbf{X}_{N})^{-1}\,\mathbf{X}_{N}^{T}\mathbf{Z}_{N}, \tag{26}$$

$$\hat{\boldsymbol{\delta}} = (\mathbf{Y}_{N}^{T}\mathbf{Y}_{N})^{-1}\,\mathbf{Y}_{N}^{T}\mathbf{Z}_{N}, \tag{27}$$

$$\hat{\mathbf{E}} = \frac{1}{N - (g - l) - 1}\,(\mathbf{Z}_{N} - \mathbf{X}_{N}\hat{\boldsymbol{\alpha}} - \mathbf{Y}_{N}\hat{\boldsymbol{\delta}})^{T} \times (\mathbf{Z}_{N} - \mathbf{X}_{N}\hat{\boldsymbol{\alpha}} - \mathbf{Y}_{N}\hat{\boldsymbol{\delta}}). \tag{28}$$

To test the entire model, the total matrix $\mathbf{T}$ of sums of squares and cross products is split into two orthogonal matrices: the hypothesis matrix, $\mathbf{H}_{g}$, and the residual matrix, $\mathbf{R}$ (Table 3). Furthermore, $\mathbf{H}_{g}$ can be orthogo-

TABLE 3

Multivariate analysis of variance table for the model given by expression 4

Source of variation	SSP matrix	d.f.	SSP matrix expectation
Over all loci	$\mathbf{H_g} = (\mathbf{X_N}\hat{\boldsymbol{\alpha}} + \mathbf{Y_N}\hat{\boldsymbol{\delta}})^T(\mathbf{X_N}\hat{\boldsymbol{\alpha}} + \mathbf{Y_N}\hat{\boldsymbol{\delta}}) - N\bar{\mathbf{z}}^{\mathbf{T}}\bar{\mathbf{z}}$	$g - l$	$(g - l)\mathbf{E} + N\mathbf{G}$
Regression allelic doses	$\mathbf{H_a} = (\mathbf{X_N}\hat{\boldsymbol{\alpha}})^T(\mathbf{X_N}\hat{\boldsymbol{\alpha}}) - N\bar{\mathbf{z}}^{\mathbf{T}}\bar{\mathbf{z}}$	$n = a - l$	$n\mathbf{E} + N\mathbf{A}$
Deviations	$\mathbf{H_d} = \mathbf{H_g} - \mathbf{H_a}$	$m = g - a$	$m\mathbf{E} + N\mathbf{D}$
Residual	$\mathbf{R} = (\mathbf{Z_N} - \mathbf{X_N}\hat{\boldsymbol{\alpha}} - \mathbf{Y_N}\hat{\boldsymbol{\delta}})^T(\mathbf{Z_N} - \mathbf{X_N}\hat{\boldsymbol{\alpha}} - \mathbf{Y_N}\hat{\boldsymbol{\delta}})$	$N - (g - l) - 1$	$[N - (g - l) - 1]\mathbf{E}$
Total	$\mathbf{T} = \mathbf{Z_N^T Z_N} - N\bar{\mathbf{z}}^{\mathbf{T}}\bar{\mathbf{z}}$	N − 1	

$\bar{\mathbf{z}}$ is a $(t \times 1)$ vector of quantitative trait means.

nally partitioned into the $\mathbf{H_a}$ and $\mathbf{H_d}$ matrices to test for additive and dominance effects associated with the markers. $\mathbf{H_a}$ and $\mathbf{H_d}$ may be obtained by fitting the $\mathbf{Z_N}$ matrix to the matrices of individual allelic doses and allelic dose products, $\mathbf{X_N}$ and $\mathbf{Y_N}$, respectively. An alternative but equivalent method to compute $\mathbf{H_g}$ is to perform a multivariate analysis of the variance (MANOVA) of the $\mathbf{Z_N}$ matrix over all loci, considering only the main effects (Table 3). The genotypes at each locus are the levels for each factor. The Wilks' Λ statistic is one of the most common test statistic computed by standard multivariate software packages. It depends on the eigenvalues of matrix $\mathbf{R}^{-1}\mathbf{H}$, where $\mathbf{H}$ is any matrix hypothesis to be tested (ANDERSON 1984).

Unbiased estimators of the matrices of marker contributions to the genetic, additive and dominance variances-covariances, **G**, **A** and **D**, can be obtained from Table 3, where we give the expectations of the matrices of sums of squares and cross products for each source of variation. The significance of each diagonal element (variance) in matrix **A** is tested by multiple regression analysis on the allelic doses (see also LANDE and THOMPSON 1990). Similar tests may be carried out for the diagonal elements in **G** and **D**. The off-diagonal elements of matrix **A** (covariances) can be tested by estimating the correlation between the linear functions i and j (LI 1975). For a given off-element of the $\mathbf{H_a}$ matrix, this correlation can be computed as follows:

$$\hat{r}_{z_i z_j} = \frac{SPH_{a,ij}}{\sqrt{SSR_{ii}SSR_{jj}}}. \qquad (29)$$

The test statistic for this correlation coefficient is

$$t = \frac{r}{\sqrt{1 - r^2}}\sqrt{N - n - 2}. \qquad (30)$$

If the null hypothesis of no covariance is true, the statistic follows a Student's distribution with N-n-2 degrees of freedom. Similar procedures can be used to test the off-diagonal elements in matrices **G** and **D**, except that the degrees of freedom are, respectively, N-(g-l)-2 and N-(g-a)-2. A lower bound for the narrow-sense heritability of each quantitative trait can be obtained by dividing each diagonal element in the **A** matrix by the same diagonal element in the **T** matrix. Finally, the specific effects of each marker allele on each quantitative trait can be tested by looking for the significance of each regression coefficient in the multiple regression analysis. When the genotype frequencies depart from Hardy-Weinberg expectations, the allelic doses and the allelic dose products are not longer uncorrelated, and the $\mathbf{H_d}$ matrix must be estimated from the difference, $\mathbf{H_g} - \mathbf{H_a}$ (Table 3). Interaction across loci, which was assumed to be absent in our model, can be tested by a MANOVA analysis over all loci, by comparing the matrix of main effects of genotypes, $\mathbf{H_g}$, with the matrix for effects of all order, $\mathbf{H_{all}}$. If $\mathbf{H_{all}} - \mathbf{H_g}$ is significant, then we can infer some interaction effect. But whether or not this effect exists, the estimation of the $\mathbf{H_g}$, $\mathbf{H_a}$ and $\mathbf{H_d}$ matrices remains unaltered.

An illustration: Data from the European oyster, *Ostrea edulis* L. (ALVAREZ *et al.* 1989), provides a practical application of our estimation procedures. The data come from a large cohort of oysters, 30 months old, located at the Ria de Ortigueira (NW Spain). Four traits were measured in each individual: shell length (SL), shell width (SW), shell depth (SD) and age-specific weight (WT). The untransformed measurements (weight in grams and shell dimensions in millimeters) were used in the analysis. For each individual, the genotype at five polymorphic allozyme loci (malate dehydrogenase, *Mdh*; phosphoglucose isomerase, *Pgi*; esterase, *Est*; isocitrate dehydrogenase, *Idh*; and phospho-glucomutase, *Pgm*) was determined by starch gel electrophoresis. Each locus was diallelic except *Pgm*, which was segregating for four alleles in this population. All statistical analyses were carried out with a sample of N = 359 individuals for which we have complete genotypic information for all five loci. No significant departure from the Hardy-Weinberg proportions was observed at any of the five loci (ALVAREZ *et al.* 1989).

The data were analyzed using SAS's GLM and REG procedures (SAS Institute Inc. 1985), which allow both MANOVA and multivariate regression analyses. We first

TABLE 4

Results of the MANOVA carried out for the data set of the European oyster

Source of variation	SSP matrix	d.f.	Wilks' lambda	P
Over all loci	$\mathbf{H_g}$	13	0.77	<0.001
Regression allelic doses	$\mathbf{H_a}$	7	0.78	<0.001
Deviations	$\mathbf{H_d}$	6	0.97	0.877
Residual	**R**	345		
Total	**T**	358		

Data from Alvarez *et al.* (1989).

TABLE 5

Matrices of estimated allele effects, α, and estimated contributions to the additive variances and covariances, A, for five allozyme loci in the European oyster

Matrix of allele effects (partial regression coefficients)

	SL	SW	SD	WT
$\hat{\alpha}$ =				
Mdh^{110}	1.05	2.71**	0.64*	1.98
Pgi^{90}	−2.49	−4.11**	0.89*	0.08
Est^{90}	4.04**	5.61***	0.94*	6.33**
Idh^{90}	1.67	3.96**	1.23***	2.87
Pgm^{80}	−5.73	−4.43	−1.93	−10.18
Pgm^{90}	−4.47	−3.28	−1.79	−7.01
Pgm^{100}	−4.84	−3.21	−1.59	−7.34

Matrix of contributions to the additive variances and covariances

	SL	SW	SD	WT
$\hat{A}$ =				
SL	1.24*	2.67	0.25	1.74
SW		5.00***	0.70	3.46
SD			0.27***	0.68
WT				2.86*

SL, shell length; SW, shell width; SD, shell depth; WT, weight. Data from Alvarez *et al.* (1989).
* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

performed a MANOVA analysis with the genotypes at all five loci as the independent variables and the measured variables as the dependent ones (Table 4). Matrix $\mathbf{H_a}$ was computed from a multivariate regression analysis where the allelic doses were used as independent variables. $\mathbf{H_d}$ was estimated as the difference $\mathbf{H_g} - \mathbf{H_a}$. Highly significant genetic effects were detected and these were exclusively additive (Table 4). All four additive variances were significant, showing a clear association between the markers and the quantitative traits (Table 5). The estimated additive contributions of the five markers, given as % of the trait phenotypic variance, were 2.15% (SL), 7.79% (SW), 5.58% (SD) and 2.88% (WT). In addition, the univariate multiple regressions showed that all loci, except *Pgm*, influenced significantly one or more traits (Table 5). Thus, we can conclude from this analysis that, at least at the moment at which this population was sampled, the four quantitative traits had non-zero heritabilities. In addition, four of the five marker loci affected the traits or were associated with regions containing QTLs. Finally, no significant interaction across loci was detected.

We also calculated the correlation of the four quantitative traits with the multilocus heterozygosity. The results showed a significant effect for three traits (all but SL). The proportion of phenotypic variance explained by this correlation (r^2) was SL 0.44% ($P = 0.21$), SW 1.64% ($P = 0.015$), SD 3.82% ($P = 0.0002$) and WT 2.13% ($P = 0.0056$). All these figures are considerably lower than the contributions to the genetic variance calculated above. Thus, as expected, the heterozygosity only captures a portion of the genetic effects: our analysis is a much more powerful way of detecting the association of phenotypic variation with marker loci.

DISCUSSION

Testing and estimating the contribution of a set of markers to the genetic variances and covariances: The computation of the correlation of phenotypic value with multilocus heterozygosity has been a method widely used in the past to study the relation between a quantitative trait and a set of markers. It has yielded positive results in some species or for some traits (SINGH and ZOUROS 1978; ZOUROS *et al.* 1980, 1988; KOEHN and GAFFNEY 1984; KOEHN *et al.* 1988; ALVAREZ *et al.* 1989; GAFFNEY 1990; POGSON and ZOUROS 1994) but negative results in others (BEAUMONT 1982; BEAUMONT *et al.* 1985; FOLZ and ZOUROS 1984; CHAKRABORTY *et al.* 1986; HOULE 1989; BOOTH *et al.* 1990). In addition, differences among loci in their contribution to the correlation have been observed in some cases (KOEHN *et al.* 1988). Our analysis, however, has shown that this is a weak way to test the relation between a set of loci and a phenotypic trait. It lacks power to detect such an effect and, even when it does give significant results, the variance explained by the correlation does not correspond in general with the variance components of known biological significance. These conclusions are true regardless whether the markers are QTLs or NMLs and are not likely to change if the assumptions of our models are relaxed, *i.e.*, by allowing for departures of the genotypic frequencies from Hardy-Weinberg proportions or for interactions between loci.

The method presented here for testing the contribution of a set of markers to the genetic variances and covariances of several quantitative traits is a multivariate

and multilocus generalization of the analysis of variance currently in use to test the contribution of one marker locus to the genetic variance of a single character (SING and DAVIGNON 1985; BOERWINKLE and SING 1986; BOERWINKLE *et al.* 1987; EDWARDS *et al.* 1987; RUIZ *et al.* 1991). In addition, we provide a way to partition, by means of the multivariate regression on the allelic doses, the contribution of the markers to the genetic variances and covariances into additive and dominance components. This partition is very important, not only because it yields variance components with a clear biological significance (LANDE 1988; FALCONER 1989) but also because it provides a more powerful test of the marker effects than the MANOVA alone. For instance, when the locus effects are exclusively additive, the regression analysis clearly increases the chances of detecting marker loci with effects on the quantitative traits relative to the MANOVA (due to the fewer degrees of freedom). That the picture obtained from the correlation with multilocus heterozygosity may change substantially when our method is applied is illustrated by a reanalysis of a data set from the European oyster (ALVAREZ *et al.* 1989). Consistent application of our method might perhaps explain the variable results of studies correlating fitness components with multilocus heterozygosity.

Interpretation of the estimated contributions: The detection of a significant effect by means of the statistical methods proposed here does not guarantee that the analyzed marker loci are actually QTLs. In some cases (*e.g.*, KOEHN *et al.* 1988; SING *et al.* 1988), additional biochemical and physiological evidence may be presented to support the hypothesis that the analyzed loci are in fact QTLs (the so called *candidate gene approach*). In many cases, however, this kind of information is lacking and the analyzed loci may be NMLs merely marking a chromosome region containing one or more QTLs. When this is the case, what kind of inferences can we make about the contribution of the QTLs from the information we have obtained from the NMLs?

The analysis carried out here shows that the additive and dominance variances explained by a set of NMLs is always lower than, or equal to, the actual additive and dominance variances contributed by the associated QTLs. This suggests that genetic markers, such as allozyme loci or RFLPs, could be used to obtain *minimum* estimates of the additive and dominance variances of quantitative traits in natural populations. This is useful in the estimation of heritabilities. This entails no special problems in the laboratory (BECKER 1984; FALCONER 1989) but is difficult in natural populations of many species, such as Drosophila, because relatives cannot be identified in the field (RISKA *et al.* 1989). Further, we might expect that a similar relation between the actual and apparent contributions to the genetic variances holds for the contributions to the genetic covariances. This, however, is not generally true: the apparent contribution of the NMLs to the genetic covariance may be lower than, equal to or larger than the total covariance contributed by the associated QTLs. This result suggests that NMLs cannot be used regularly to estimate genetic covariances between traits, and that, when an apparent contribution of a set of NMLs to the genetic covariance between two traits is detected, the results must be interpreted with caution.

A brief mention should be made of the possibility, implicit in our models, that some of the marker loci included in a given analysis are QTLs whereas the remaining loci are NMLs. All the conclusions stated in the preceding paragraph concerning the relation between the apparent and actual contributions to the genetic variances and covariances also hold true here. Furthermore, the partial regression coefficients will allow us to identify, when a NML is associated to a QTL and both are included in the analysis, which is the gene actually affecting the characters. Likewise, when several NMLs are associated with a hidden QTL (*i.e.*, one not included among the analyzed markers), the partial regression coefficients will point to the NML with the highest correlation with the QTL. This NML will usually be (see below) the marker physically closest to the hidden QTL. Therefore, the multiple regression approach allows us to identify either true QTLs or chromosome regions containing one or more QTLs. This may be an important step towards the isolation and cloning of QTLs.

The linkage disequilibrium condition: The *sine qua non* for observing an apparent contribution of the NMLs to the genetic variances and covariances of quantitative traits is, of course, linkage disequilibrium between the NMLs and the QTLs affecting the traits. In natural populations, several forces can generate disequilibria between NMLs and QTLs: random genetic drift (HILL and ROBERTSON 1968; SVED and FELDMAN 1973; HILL 1976; SLATKIN 1994), founder effect (KIMURA and OHTA 1971), population subdivision and migration (NEI and LI 1973; FELDMAN and CHRISTIANSEN 1975; CHAKRABORTY and WEISS 1988), hybridization (LANDE and THOMPSON 1990), hitchhiking (THOMSON 1977; ROBINSON *et al.* 1991b) and epistatic selection (LEWONTIN and KOJIMA 1960; BODMER and FELSENSTEIN 1967; KARLIN 1975). These forces are opposed by recombination that tend to erode the associations. Occasional hybridization between genetically differentiated lines is probably the most powerful mechanism for generating associations in domestic species (LANDE and THOMPSON 1990) and population admixture has probably been very important in humans (CHAKRABORTY and WEISS 1988). In natural populations of many species, however, random drift, migration or selection might be perhaps more relevant. SLATKIN (1994) has shown by

simulation that drift is very likely to generate significant nonrandom associations between closely linked polymorphic genes in stable populations.

Turning to the empirical data, it is clear that gametic associations are not an universal feature of all genes and all populations, but the opposite statement, namely, that there are no linkage disequilibria in natural populations, is also untrue. There are many examples of gametic associations in the literature (HEDRICK *et al.* 1978; BARKER 1979). Additionally, many cases where the authors failed to detect statistically significant linkage disequilibrium may merely reflect the low power of the tests used (ZAPATA and ALVAREZ 1992, 1993). As expected disequilibria are more intense as recombination decreases, *i.e.*, between closely linked markers (LANGLEY 1977; AQUADRO *et al.* 1986; ZAPATA and ALVAREZ 1992, 1993). Perhaps the best, yet indirect, evidence for pervasive genome-wide linkage disequilibrium is the positive correlation between nucleotide diversity and recombination rate, which has been explained in terms of hitchhiking with favorable mutations (BEGUN and AQUADRO 1992). Disequilibria are also strong is some special cases, for example, when the loci are linked to chromosome inversions, which are abundant in Drosophila and other insects (KRIMBAS and POWELL 1992). Finally, gametic disequilibria are present *within* families even in random mating populations. This fact has been used for a long time to detect QTLs in man and domestic animals (HALEY 1991; KNOTT and HALEY 1992).

In any case, whatever the cause, empirical observations show that markers with an effect on quantitative traits are found surprisingly frequently. For instance, SING and ORR (1976) examined the associations between 12 unselected red cell and serum markers and variation in total serum cholesterol levels. They found that one third of the analyzed markers accounted for statistically significant portions of the variance in this quantitative trait. Also, in the studies cited above on the correlation between growth rate and multilocus heterozygosity, the authors typically use only a small number of allozyme markers (6 on the average) and, in many cases, in spite of the low number of markers used, an effect is detected with the correlation explaining ~5–10% of the phenotypic variation of the trait. KOEHN *et al.* (1988), who analyzed an unusually large number of polymorphic loci (15) in the coot clam *Mulinia lateralis,* found that 8 of them showed a significant effect on growth rate. In the example analyzed here, four of the five markers showed a significant effect upon two or more of the quantitative traits. These examples suffice to show that estimating the contributions to the additive and dominance variances using markers is practical; it is by no means a grail-like pursuit.

A draft of this paper was written while the first author was Visiting Scholar in the Department of Ecology and Evolutionary Biology of the University of Arizona (Tucson, AZ). He is indebted to BILL and SARAH HEED, HENAR ALONSO-PIMENTEL, KIYOMI MORINO and GREG HUCKINS for their warm hospitality and unfailing support during his stay in Tucson. We are specially grateful to GONZALO ALVAREZ and CARLOS ZAPATA (Universidad de Santiago de Compostela), who made available to us the data of the European oyster used in the illustration. We also thank BRUCE WALSH for his constructive comments on the manuscript and for sharing with us unpublished materials. ANDREW BERRY, RUSSELL LANDE, DICK LEWONTIN, ARCADI NAVARRO, ESTHER BETRÁN, JOSÉ M. RANZ, CARLOS ZAPATA and an anonymous referee also made many useful comments. Work was supported by grant PB93–0844 from the DGICYT (Ministerio de Educación y Ciencia, Spain) to A.R. and a postdoctoral MEC/Fulbright fellowship from the Ministerio de Educación y Ciencia (Spain) to A.B.

LITERATURE CITED

ABLER, B. S. B., M. D. EDWARDS and C. W. STUBER, 1991 Isoenzymatic identification of quantitative trait loci in crosses of elite maize inbreds. Crop Sci. **31:** 267–274.

ALVAREZ, G., C. ZAPATA, R. AMARO and A. GUERRA, 1989 Multilocus heterozygosity at protein loci and fitness in the european oyster, *Ostrea edulis* L. Heredity **63:** 359–372.

ANDERSON, T. W., 1984 *An Introduction to Multivariate Analysis,* Ed. 2, J. Wiley, New York.

AQUADRO, C. F., S. F. DESSE, M. M. BLAND, C. H. LANGLEY and C. C. LAURIE-AHLBERG, 1986 Molecular population genetics of the alcohol dehydrogenase gene region of *Drosophila melanogaster.* Genetics **114:** 1165–1190.

AVERY, P. J., and W. G. HILL, 1977 Variability in genetic parameters among small populations. Genet. Res. **29:** 193–213.

AVERY, P. J., and W. G. HILL, 1979 Variance in quantitative traits due to linked dominant genes and variance in heterozygosity in small populations. Genetics **91:** 817–844.

BARKER, J. S. F., 1979 Inter-locus interactions: a review of experimental evidence. Theor. Popul. Biol. **16:** 323–346.

BEAUMONT, A. R., 1982 Variations in heterozygosity at two loci between year classes of a population of *Chlamys opercularis* (L.) from a Scottish sea-loch. Mar. Biol. Lett. **3:** 24–34.

BEAUMONT, A. R., E. M. GOSLING, C. M. BEVERIDGE, M. D. BUDD and G. M. BURNELL, 1985 Studies on heterozygosity and size in the scallop, *Pecten maximus,* pp. 443–454 in *Proceedings of the 19th European Marine Biology Symposium,* edited by P. E. GIBBS. Cambridge University Press, Cambridge.

BECKER, W. A., 1984 *Manual of Quantitative Genetics,* Ed. 4, Academic Enterprises, Pullman.

BEGUN, D. J., and C. F. AQUADRO, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster.* Nature **356:** 519–520.

BODMER, W. F., and J. FELSENSTEIN, 1967 Linkage and selection: theoretical analysis of the deterministic two-locus random mating model. Genetics **57:** 237–265.

BOERWINKLE, E., and C. F. SING, 1986 Bias of the contribution of single locus effects to the variance of a quantitative trait. Am. J. Hum. Genet. **39:** 137–144.

BOERWINKLE, E., S. VISVIKIS, D. WELSH, J. STEINMETZ, S. M. HANASH *et al.,* 1987 The use of measured genotype information in the analysis of quantitative phenotypes in man. II. The role of the apolipoprotein E polymorphism in determining levels, variability and covariability of cholesterol, betalipoprotein, and triglycerides in a sample of unrelated individuals. Am. J. Med. Genet. **27:** 567–582.

BOOTH, C. L., D. S. WOODRUFF and S. J. GOULD, 1990 Lack of significant associations between allozyme heterozygosity and phenotypic traits in the land snail *Cerion.* Evolution **44:** 210–213.

CHAKRABORTY, R., and N. RYMAN, 1983 Relationship of mean and variance of genotypic values with heterozygosity per individual in a natural population. Genetics **103:** 149–152.

CHAKRABORTY, R., and K. M. WEISS, 1988 Admixture as a tool for

finding linked genes and detecting that difference from allelic association between loci. Proc. Natl. Acad. Sci. USA **85:** 9119–9123.

CHAKRABORTY, R., R. E. FERRELL, S. A. BARTON and W. J. SCHULL, 1986 Genetic polymorphism and fertility parameters in the Aymara of Chile and Bolivia. Ann. Hum. Genet. **50:** 69–82.

COCKERHAM, C. C., 1954 An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. Genetics **39:** 859–882.

DEVICENTE, M. C., and S. D. TANKSLEY, 1993 QTL Analysis of transgressive segregation in an interspecific tomato cross. Genetics **134:** 585–596.

EDWARDS, M. D., C. W. STUBER and J. F. WENDEL, 1987 Molecular marker facilitated investigations of quantitative trait loci in maize. I. Numbers, distribution, and types of gene action. Genetics **116:** 113–125.

EWENS, W. J., 1979 *Mathematical Population Genetics.* Springer-Verlag, Berlin.

FALCONER, D. S., 1989 *Introduction to Quantitative Genetics,* Ed. 3, Longman, London.

FELDMAN, M. W., and F. B. CHRISTIANSEN, 1975 The effect of population subdivision on two loci without selection. Genet. Res. **24:** 151–162.

FOLZ, D. W., and E. ZOUROS, 1984 Enzyme heterozygosity in the scallop *Placopecten magellanicus* in relation to age and size. Mar. Biol. Lett. **5:** 255–263.

GAFFNEY, P. M., 1990 Enzyme heterozygosity, growth rate, and viability in *Mytilus edulis*: another look. Evolution **44:** 204–210.

HALEY, C. S., 1991 Use of DNA fingerprints for the detection of major genes for quantitative traits in domestic species. Anim. Genet. **22:** 259–277.

HALEY, C. S., and S. A. KNOTT, 1992 A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity **69:** 315–324.

HEDRICK, P., S. JAIN and L. HOLDEN, 1978 Multilocus systems in evolution. Evol. Biol. **11:** 101–184.

HILL, W. C., 1976 Non-random association of neutral linked genes in finite populations, pp. 339–376 in *Population Genetics and Ecology,* edited by S. KARLIN and E. NEVO. Academic Press, New York.

HILL, W. C., and A. ROBERTSON, 1968 Linkage disequilibrium in finite populations. Theor. Appl. Genet. **38:** 226–231.

HOULE, D., 1989 Allozyme-associated heterosis in *Drosophila melanogaster.* Genetics **123:** 789–801.

KARLIN, S., 1975 General two-locus selection models: some objectives, results and interpretations. Theor. Popul. Biol. **7:** 364–398.

KIMURA, M., and T. OHTA, 1971 *Theoretical Aspects of Population Genetics.* Princeton University Press, Princeton, NJ.

KNAPP, S. J., W. C. BRIDGES, JR. and D. BIRKES, 1990 Mapping quantitative trait loci using molecular marker linkage maps. Theor. Appl. Genet. **79:** 583–592.

KNOTT, S. A., and C. S. HALEY, 1992 Maximum likelihood mapping of quantitative trait loci using full sib families. Genetics **132:** 1211–1222.

KOEHN, R. K., and P. M. GAFFNEY, 1984 Genetic heterozygosity and growth rate in *Mytilus edulis.* Mar. Biol. **82:** 1–7.

KOEHN, R. K., W. J. DIEHL and T. M. SCOTT, 1988 The differential contribution of individual enzymes of glycolysis and protein catabolism to the relationship between heterozygosity and growth rate in the coot clam *Mulinia lateralis.* Genetics **118:** 121–130.

KRIMBAS, C. B., and J. R. POWELL (Editors), 1992 *Drosophila Inversion Polymorphism.* CRC Press, London.

LANDE, R., 1988 Quantitative genetics and evolutionary theory, pp. 71–84 in *Proceedings of the Second International Conference on Quantitative Genetics,* edited by B. S. WEIR, M. M. GOODMAN, E. J. EISEN and G. NAMKOONG. Sinauer, Sunderland, MA.

LANDE, R., and R. THOMPSON, 1990 Efficiency of marked-assisted selection in the improvement of quantitative traits. Genetics **124:** 743–756.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

LANGLEY, C. H., 1977 Nonrandom associations between allozymes in natural populations of *Drosophila melanogaster,* pp. 265–273 in *Lecture Notes in Biomathematics,* Vol. 19, *Measuring Selection in Natural Populations,* edited by F. B. CHRISTIANSEN and T. M. FENCHEL. Springer-Verlag, New York.

LEWONTIN, R. C., and K. KOJIMA, 1960 The evolutionary dynamics of complex polymorphisms. Evolution **14:** 458–472.

LI, C. C., 1975 *Path Analysis—A Primer.* The Boxwood Press, Pacific Grove, CA.

MARTINEZ, O., and R. N. CURNOV, 1992 Estimating the locations and sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet. **85:** 480–488.

MITTON, J. B., and M. C. GRANT, 1984 Associations among protein heterozygosity, growth rate and developmental homeostasis. Annu. Rev. Ecol. Syst. **15:** 479–499.

MORENO-GONZALEZ, J., 1992 Genetic models to estimate additive and non-additive effects of marker-associated QTL using multiple regression techniques. Theor. Appl. Genet. **85:** 435–444.

NEI, M., and W. H. LI, 1973 Linkage disequilibrium in subdivided populations. Genetics **75:** 213–219.

NODARI, R. O., S. M. TSAI, P. GUZMAN, R. L. GILBERTSON and P. GEPTS, 1993 Toward an integrated linkage map of common bean. III. Mapping genetic factors controlling host-bacteria interactions. Genetics **134:** 341–350.

PATERSON, A. H., S. DAMON, J. D. HEWITT, D. ZAMIR, H. D. RABINOWITCH *et al.*, 1991 Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. Genetics **127:** 181–197.

PATERSON, A. H., E. S. LANDER, J. D. HEWITT, S. PATERSON, S. E. LINCOLN *et al.*, 1988 Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature **335:** 721–726.

POGSON, G. H., and E. ZOUROS, 1994 Allozyme and RFLP heterozygosities as correlates of growth rate in the scallop *Placeopecten magellanicus*: a test of the associative overdominance hypothesis. Genetics **137:** 221–231.

RISKA, B., T. PROUT and M. TURELLI, 1989 Laboratory estimates of heritabilities and genetic correlations in nature. Genetics **123:** 865–871.

ROBINSON, W. P., M. A. ASMUSSEN and G. THOMPSON, 1991a Three-locus systems impose additional constrains on pairwise disequilibria. Genetics **129:** 925–930.

ROBINSON, W. P., A. CAMBON-THOMSEN, N. BOROT, W. KLITZ, and G. THOMSON, 1991b Selection, hitchhiking and disequilibrium analysis at three linked loci with application to HLA data. Genetics **129:** 931–948.

RUIZ, A., M. SANTOS, A. BARBADILLA, J. E. QUEZADA-DIAZ, E. HASSON *et al.*, 1991 Genetic variance for body size in a natural population of *Drosophila buzzatii.* Genetics **128:** 739–750.

SAX, K., 1923 The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris.* Genetics **8:** 552–560.

SING, C. F., and J. DAVIGNON, 1985 Role of the apolipoprotein polymorphism in determining normal plasma lipid and lipoprotein variation. Am. J. Hum. Genet. **37:** 268–285.

SING, C. F., and J. ORR, 1976 Analysis of genetic and environmental sources of variation in serum cholesterol in Tecumseh, Michigan. III. Identification of genetic effects using twelve polymorphic blood marker systems. Am. J. Hum. Genet. **28:** 453–464.

SING, C. F., E. BOERWINKLE, P. P. MOLL and A. R. TEMPLETON, 1988 Characterization of genes affecting quantitative traits in humans, pp. 250–269 in *Proceedings of the Second International Conference on Quantitative Genetics,* edited by B. S. WEIR, M. M. GOODMAN, E. J. EISEN and G. NAMKOONG. Sinauer, Sunderland, MA.

SINGH, S. M., and E. ZOUROS, 1978 Genetic variation associated with growth rate in the American oyster (*Crassostrea virginica*). Evolution **32:** 342–352.

SLATKIN, M., 1994 Linkage disequilibrium in growing and stable populations. Genetics **137:** 331–336.

STUBER, C. W., M. D. EDWARDS and J. F. WENDEL, 1987 Molecular marker-facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. Crop Sci. **27:** 639–648.

STUBER, C. W., S. E. LINCOLN, D. W. WOLFF, T. HELENTJARIS, and E. S. LANDER, 1992 Identification of genetic factors contributing

to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics **132:** 823–839.

SVED, J. A., and M. W. FELDMAN, 1973 Correlation and probability methods for one and two loci. Theor. Popul. Biol. **4:** 129–132.

TANKSLEY, S. D., H. MEDINA-FILHO and C. M. RICK, 1982 Use of naturally occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity **49:** 11–25.

THODAY, J. M., 1961 Location of polygenes. Nature **191:** 368–370.

THOMSON, G., 1977 The effect of a selected locus on linked neutral loci. Genetics **85:** 753–788.

VOGLER, G. P., and D. W. FULKER, 1988 Human behavior genetics, pp. 475–503 in *Handbook of Multivariate Experimental Psychology*, edited by J. NESSELROAD and R. B. CATTELL. Plenum Press, NY.

WEIR, B. S., 1990 *Genetic Data Analysis.* Sinauer, Sunderland, MA.

WEIR, B. S., and C. C. COCKERHAM, 1977 Two-locus theory in quantitative genetics, pp. 247–269 in *Proceedings of the International Conference on Quantitative Genetics*, edited by E. POLLAK, O. KEMPTHORNE and T. B. BAILEY, JR. Iowa State Univ. Press, Ames.

WEIR, B. S., and C. C. COCKERHAM, 1989 Complete characterization of disequilibrium at two loci, pp. 86–110 in *Mathematical Evolutionary Theory*, edited by M. E. FELDMAN. Princeton University Press, Princeton, NJ.

WEIR, B. S., C. C. COCKERHAM and J. REYNOLDS, 1980 The effects of linkage and linkage disequilibrium on the covariances of noninbred relatives. Heredity **45:** 351–359.

WELLER, J. I., 1986 Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. Biometrics **42:** 627–640.

WELLER, J. I., 1987 Mapping and analysis of quantitative trait loci in Lycopersicon (tomato) with the aid of genetic markers using approximate maximum likelihood methods. Heredity **59:** 413–421.

WRIGHT, S., 1921 Correlation and causation. J. Agric. Res. **20:** 557–585.

WRIGHT, S., 1934 The method of path coefficients. Ann. Math. Stat. **5:** 161–215.

WRIGHT, S., 1968 *Evolution and the Genetics of Populations.* Vol. I. *Genetic and Biometric Foundations.* University of Chicago Press, Chicago.

ZAPATA, C., and G. ALVAREZ, 1992 The detection of gametic disequilibrium between allozyme loci in natural populations of Drosophila. Evolution **46:** 1900–1917.

ZAPATA, C., and G. ALVAREZ, 1993 On the detection of nonrandom associations between DNA polymorphisms in natural populations of Drosophila. Mol. Biol. Evol. **10:** 823–841.

ZOUROS, E., and D. W. FOLTZ, 1987 The use of allelic isozyme variation for the study of heterosis. Curr. Top. Biol. Med. Res. **15:** 1–60.

ZOUROS, E., M. ROMERO-DOREY and A. L. MALLET, 1988 Heterozygosity and growth in marine bivalves: further data and possible explanations. Evolution **42:** 1332–1341.

ZOUROS, E., S. M. SINGH and H. E. MILES, 1980 Growth rate in oysters: an overdominant phenotype and its possible explanations. Evolution **34:** 856–867.

Communicating editor: B. S. WEIR