Differentiation of Allelic Frequencies at Quantitative Trait Loci Affecting Locally Adaptive Traits

Robert G. Latta*

Department of Environmental, Population and Organismic Biology, University of Colorado, Boulder, Colorado 80309-0334

Submitted May 5, 1997; Accepted September 10, 1997

Keywords: quantitative trait loci, linkage disequilibrium, multilocus differentiation, local adaptation.

Direct comparisons of population differentiation observed in neutral loci with that seen in quantitative traits have revealed discordant patterns, most commonly that quantitative traits are more differentiated than single neutral loci. For example, body size has a greater proportion of genetic variation attributable to among-population differences than do presumed neutral allozymes in both *Drosophila buzzatii* (Prout and Barker 1993) and *Daphnia obtusa* (Spitze 1993). Podolsky and Holtsford (1995) examined a suite of discrete and continuous traits in *Clarkia dudleyana* and found that six of the 18 traits were more differentiated than allozyme markers. Four of six morphological traits studied by Yang et al. (1996) in lodgepole pine (*Pinus contorta*) showed greater differentiation than neutral allozymes. Each of these studies provides evidence that quantitatively inherited traits are experiencing diversifying selection which acts to adapt populations to local conditions, a conclusion that is strengthened by reciprocal transplant experiments showing adaptation of populations to their local environments (e.g., Schemske 1984; Jordan 1991). By contrast, some traits show less differentiation than do neutral markers (e.g., Spitze 1993; Bonnin et al. 1996), suggesting that stabilizing selection across environments acts to resist differentiation of these traits.

Such comparisons have proceeded by testing the neutral theory prediction that observed within- and between-population components of variance in quantitative traits are consistent with observed differentiation at neutral loci, F_{st}. If additive variation in a quantitative trait differentiates among populations solely by the neutral processes of migration and drift, then the genetic variances for the trait will be

$$\sigma_b^2 = 2F_{st}\sigma_0^2,$$
$$\sigma_w^2 = (1 - F_{st})\sigma_0^2, \quad (1)$$

and

$$\sigma_t^2 = (1 + F_{st})\sigma_0^2,$$

where σ_w^2, σ_b^2, and σ_t^2 are the within-population, between-population, and total genetic variances in the trait, respectively, and σ_0^2 is the genetic variance for the trait expected if the populations formed a single panmictic unit (Wright 1951, 1952; Lande 1992). Thus an estimate of population differentiation for quantitative traits (termed Q_{st} by Spitze 1993) can be computed as

$$Q_{st} = \frac{\sigma_b^2}{\sigma_b^2 + 2\sigma_w^2} \quad (2)$$

and compared with the level of differentiation of neutral loci, F_{st}. Some authors (e.g., Lascoux et al. 1996) have interpreted Q_{st} as the value of F_{st} for allelic variation at quantitative trait loci (QTLs) underlying the trait. However, this is true only if there is linkage equilibrium among QTLs (Rogers and Harpending 1983), an assumption that is only reasonable if the trait and all QTLs are selectively neutral.

Given that quantitative traits often show higher levels of differentiation than electrophoretic loci do, and are therefore presumably not neutral, it is relevant to ask what this implies for allelic frequencies at loci underlying locally adaptive quantitative traits. In addition, there has been some interest recently in multilocus measures of differentiation and their possible relationship to quantitative traits (e.g., Lagercrantz and Ryman 1990; Kremer et al. 1997). In this study, I will model the patterns of allelic differentiation and disequilibria expected for QTLs underlying locally adaptive traits. The results show that local selection on the trait will affect the covariance of allelic frequencies among loci (i.e., the among-population component of linkage disequilibrium; Ohta 1982) to a

*Present address: Department of Biology, McGill University, 1205 Avenue Dr. Penfield, Montreal, Quebec H3A 1B1, Canada; E-mail: rlatta@bio1.lan.mcgill.ca.

Am. Nat. 1998. Vol. 151, pp. 283–292.

much greater extent than the allelic frequencies themselves, such that individual QTLs will show levels of differentiation similar to those expected for neutral loci.

The Model

Following Falconer (1989), consider a trait, z, influenced by n loci, each with two alleles. At the i^{th} locus, one allele increases z by an amount α_i (the "+" allele), the other decreases z by the same amount. If effects are additive across loci, then the genetic variance exhibited by the trait is:

$$\sigma_z^2 = \sum_i \sigma_i^2 + \sum_i \sum_{j \neq i} \text{cov}_{(i,j)}, \tag{3}$$

where σ_i^2 is the variance contributed by allelic variation at locus i, and $\text{cov}_{(i,j)}$ is the covariance of allelic effects at loci i and j (through linkage disequilibrium). For the two allele model, σ_i^2 is $2p_iq_i\alpha_i^2$, for allelic frequencies $p_i + q_i = 1$. The covariance of allelic effects is $4D_{ij}\alpha_i\alpha_j$, where D_{ij} is the linkage disequilibrium between the loci. Equation (3) can be rewritten in matrix notation as

$$\sigma_z^2 = \mathbf{A}^{\mathrm{T}}\mathbf{C}\,\mathbf{A}, \tag{4}$$

where $\mathbf{A}$ is a column matrix of the allelic effects at each locus, α_i, and $\mathbf{C}$ is the variance-covariance matrix of genotypes across loci. Matrix $\mathbf{C}$ has diagonal elements equal to the genetic variance (expected heterozygosity, H_i) and off-diagonal elements of $4D_{ij}$.

Differentiation in the trait, z, is measured by partitioning the total variance into between- and within-population components (eq. [1]). We can partition $\mathbf{C}$ in the same manner ($\mathbf{C}_t = \mathbf{C}_b + \mathbf{C}_w$), following Ohta (1982). The diagonal elements of these matrices are the genetic variances:

$$\begin{aligned} H_{t(i)} &= 2\overline{p_i}\,\overline{q_i} \qquad H_{w(i)} = \overline{2p_iq_i} \\ H_{b(i)} &= 2\overline{p_i}\,\overline{q_i} - \overline{2p_iq_i}, \end{aligned} \tag{5a}$$

where overbars indicate averages over populations, while the off-diagonal elements are linkage disequilibria, where p_{ij} is the frequency of gametes carrying the "+" allele at loci i and j:

$$\begin{aligned} D_{t(ij)} &= \overline{p_{ij}} - \overline{p_i}\,\overline{p_j} \qquad D_{w(ij)} = \overline{p_{ij} - p_ip_j} \\ D_{b(ij)} &= \overline{p_ip_j} - \overline{p_i}\,\overline{p_j}. \end{aligned} \tag{5b}$$

The $D_{b(ij)}$ variable represents the covariance of allelic frequencies across populations. The individual values of $D_{b(ij)}$ for any pair of loci can take many values, but their aggregate effect on trait differentiation, $\sum\sum D_{b(ij)}$, can be predicted by rearranging equation (3). Assume that within each population stabilizing selection acts to move the trait toward a locally optimal value, denoted z_x^* for population x. At equilibrium, the mean trait values in each population will approach these optimal values, assuming that selection is strong enough to overcome gene flow. Thus the between-population variance in trait values σ_b^2, will approach the variance of optima ($\sigma_b^2 \to \sigma_{z^*}^2$). The between-population variance expected due to drift ($\sum\sigma_i^2$) will approach $2F_{st}\,\sigma_0^2$ from equation (1). Thus the covariance of QTL allelic frequencies across loci will contribute to the differentiation of the trait an amount approximately equal to the difference between the two:

$$\sum_i \sum_{j \neq i} \text{cov}_{(i,j)} = \sigma_{z^*}^2 - 2F_{st}\sigma_0^2. \tag{6}$$

Thus with uniform selection, such that all populations are selected to similar optima (low $\sigma_{z^*}^2$) and limited migration among populations ($2F_{st}\,\sigma_0^2$ is high), negative covariances of allelic frequencies will result. This is illustrated in figure 1*A*. Changes in the frequency of the + allele at locus 1 are counterbalanced by opposite changes at locus 2 such that there is little variance in trait values despite pronounced differentiation of allele frequencies at the two loci. Conversely, with strong diversifying selection (high $\sigma_{z^*}^2$) and extensive gene flow ($2F_{st}\,\sigma_0^2$ is low), allelic frequencies at QTLs will be positively correlated across environments. Parallel changes in allelic frequency across loci (fig. 1*B*) will reinforce each other, such that strong differentiation of trait values is possible with little differentiation of allelic frequencies. Thus the mean value of the trait in each population can potentially evolve to the optimum for its environment without the allelic frequencies at the underlying QTLs being any more or less differentiated than expected by migration/drift equilibrium.

One condition must be placed on this result. The covariance between two variables cannot exceed the larger of the variances (i.e., covariance cannot occur without variance). Thus, for n loci, at least $1/n$ of the between-population variance must be attributable to differentiation of allelic frequencies. When the difference between the variance of trait optima, ($\sigma_{z^*}^2$), and the neutral expected variance ($2F_{st}\,\sigma_0^2$) is greater than the maximum covariance of allelic frequencies, then the allelic frequencies of QTLs must become more differentiated than predicted by migration/drift balance, if populations are to adapt to local conditions. No such limitation exists for the case where uniform selection acts with low migration because negative covariances among loci can act to completely eliminate variance in the trait means (fig. 1*A*).

Computer Simulation

I used computer simulation to assess the patterns of allelic covariances across populations under local selection and, thus, test the predictions of equation (6). The simulation model assumed an array of 10 populations with an

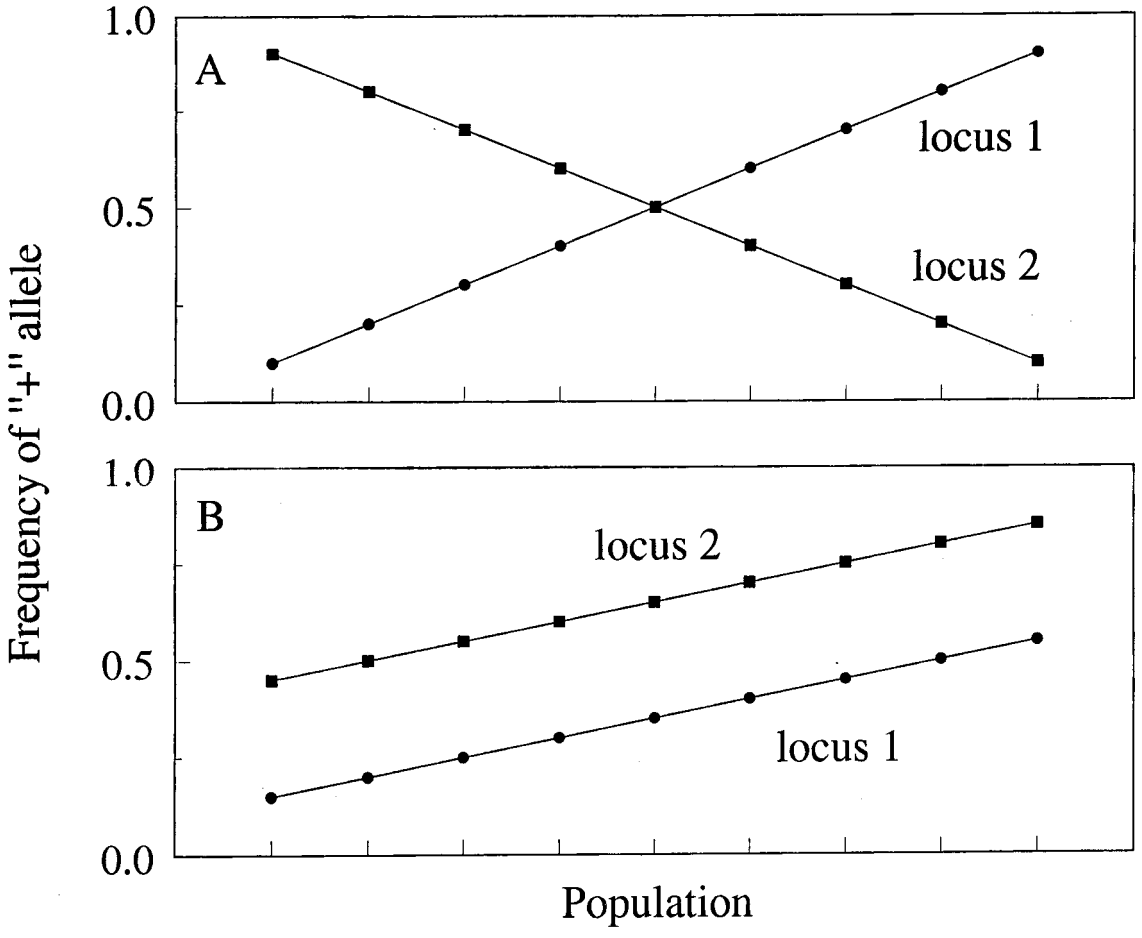

Figure 1: Hypothetical pattern of QTL allelic frequencies across populations under different conditions of selection and migration. *A,* Low variance of trait optima and low migration: negative covariance frequency of the "+" allele at locus 1 and locus 2. *B,* High variance of trait and high gene flow: positive covariance of allele frequencies.

island migration model. I further assumed that all individuals in a population mated at random and experienced uniform selection toward the local optimum, z_x^*, which varied among populations. The trait was controlled by either five loci with major effects, typical of the results from many QTL mapping studies (e.g., Paterson et al. 1991), or 20 loci each with minor effects. Loci were unlinked.

The model contained three steps: migration, mating, and selection, in that order. Allelic frequencies after migration were

$$p'_{i,x} = \sum_y p_{i,y} m_{x,y},$$

where m_{xy} is the migration rate between population x and y (m_{xx} is the frequency of nonmigrants). Within populations, 200 diploid genotypes were created by randomly drawing alleles from the array of allele frequencies for that population. Each individual's trait value was calculated as the sum of the allelic values across all loci. Selection was then imposed on the trait values as

$$w_j = e^{(-(z_j - z_x^*)^2/2\sigma_s^2)},$$

where z_x^* is the optimal value of z in population x, w_j and z_j are the fitness and trait value, respectively, of individual j, and σ_s^2 is the strength of selection on genotypes, defined by Turelli (1984). The strength of selection on genotypes, σ_s^2, is the sum of σ_e^2, the environmental variance for the trait, and σ_{fit}^2, the variance (width) of the Gaussian fitness curve. Since a given intensity of selection can be achieved with either strong selection and high environmental variance or weak selection and low environmental variance, I specified σ_s^2, rather than both σ_{fit}^2 and σ_e^2. The allelic frequencies after selection were then

$$p''_{i,x} = \frac{\left(\sum_j w_j n_{i,j}\right)}{\overline{W}_x},$$

where $n_{i,j}$ is the number of "+" alleles at locus i carried by individual j (0, 1, or 2) and $\overline{W}_x$ is the mean fitness in population x.

Parameters

The number of migrants per generation, Nm, took values of 0.1, 1, and 20 individuals per generation. These values are representative of the range inferred from allozyme loci in plants (e.g., Hamrick and Godt 1990). Low, medium, and high variance in optimal trait values were specified so as to equal the neutral expected variance for 20, 1, and 0.1 migrants, respectively. In this way the predictions of equation (6) could be tested. For example, if the variance of the trait optima corresponds to the neutral expectation for 20 migrants per generation, but migration rates are lower than this, negative covariances of allelic frequency across populations are predicted (eq. [6]). In addition, a neutral case (no selection) was modeled, along with a case of "extreme" diversifying selection, in which trait optima were set to evenly span the range of trait values that were genetically possible. Selection intensities were specified by the mean fitness of genotypes expected for a population in which $p = q = 0.5$ at all QTLs (the maximum genetic variance possible). Mean fitness is related to σ_s^2 and the genetic variance σ_g^2 as

$$\overline{W} = \sqrt{\frac{\sigma_s^2}{\sigma_s^2 + \sigma_g^2}}.$$

I examined the case of moderate selection ($\overline{W}_x = 0.975$). This value corresponds to mean fitness of phenotypes (where environmental variance reduces mean fitness) of 0.7–0.9 (assuming trait heritabilities between 0.1 and 0.25). The observed mean fitness for weakly selected traits in *Drosophila* is 0.91–0.98 (Turelli 1984), while Bradshaw (1984) has shown that fitness of immigrants to plant populations can be as low as 0.5 or less. Thus the value used here represents a moderate strength of selection.

For each combination of $\sigma_{z^*}^2$ and Nm, I ran 40 replicates of the simulation for 500 generations. This was long enough for the simulation output to reach a stationary state (i.e., results from 500 generations fell within the range over which results fluctuated for the subsequent 1,000 generations). At the end of each run, F_{st} was recorded for all loci affecting the trait. In addition, five neutral loci were included in each run to give an estimate of differentiation expected at neutral electrophoretic markers unlinked to QTLs. For each run, the proportion of trait variance attributable to differentiation among populations, σ_b^2/σ_t^2, was recorded. The proportion of σ_b^2, attributable to differentiation (variance) of allelic frequencies was quantified as $\sum\sigma_{(i)}^2/\sigma_b^2$ (the first term of eq. [3] over the total). The remainder of σ_b^2 derives from covariances. The term $\sum\sigma_{(i)}^2/\sigma_b^2$, can take values from $1/n$ (see above) to infinity, where values <1 indicate that allelic differentiation is less than trait differentiation, and positive covariances among loci are present. Conversely, values >1 indicate negative covariances among loci, such that allelic differentiation exceeds trait differentiation.

Results

In the case where variation in z was neutral, differentiation was roughly equal whether measured by F_{st} of neutral loci, F_{st} of QTLs, or Q_{st} of quantitative trait variation (table 1). Thus, the ratio of between-population to total variance in neutral trait values was consistent with equation (1). For such neutral traits, allelic frequencies vary independently at each locus such that covariances of allelic frequencies across populations were zero and could be accounted for solely by the allelic differentiation (table 1). However, this was not the case when traits were subject to selection. In general, population differentiation (σ_b^2/σ_t^2) for locally adapted traits was poorly predicted from the F_{st} value of individual loci. This is true even when equation (1) is applied to QTLs, rather than to neutral electrophoretic marker loci. Conversely, Q_{st} estimated from the differentiation of the trait generally differed both from F_{st} of neutral loci and from F_{st} of the QTLs, especially when many loci affected the trait.

In most runs, F_{st} of QTL loci fell near that of the neutral markers (table 1), such that the allelic frequencies at QTLs were only slightly more differentiated than at neutral loci. For example, with five loci, high migration, and extreme diversifying selection, F_{st} for QTLs was generally about 0.042, comparable to the value of 0.007 observed at neutral loci. If there were no covariances among allelic frequencies across populations, equation (1) would predict σ_b^2/σ_t^2 to be approximately 0.081. The observed values of σ_b^2/σ_t^2 in these runs were much higher, about 0.259, giving an estimate of Q_{st} of about 0.149. Thus strong differentiation of trait values is possible with only modest differentiation of allelic frequencies at QTLs. This discrepancy is accounted for by the fact that only about 26% of σ_b^2 was attributable to differences in allelic frequencies across populations (table 1), with the remaining 74% of σ_b^2 deriving from covariances.

Even with relatively modest variance in optimal trait values, a significant proportion of the trait variance among populations was attributable to covariances of QTL allelic frequencies across populations. The strength and sign of the covariances were generally in agreement with the predictions of equation (6). When the variance

Table 1: Patterns of differentiation at loci underlying quantitative traits subject to local selection

Number of loci, selection, and Nm	F_{st} (*mean over loci*)		Q_{st}	σ_b^2/σ_t^2		$\frac{\sum \sigma_t^2}{\sigma_b^2}$
	Neutral	*QTLs*		*Observed*	*Predicted*	
5:						
N:						
.1	.574	.594	.597	.748	.745	.965
1	.195	.185	.158	.274	.312	1.066
20	.008	.007	.008	.017	.014	1.102
U:						
.1	.576	.731	.127	.226	.844	15.130
1	.178	.243	.039	.059	.391	7.040
20	.007	.007	.007	.015	.014	1.170
M:						
.1	.591	.761	.385	.556	.864	5.050
1	.175	.262	.137	.241	.415	1.980
20	.007	.008	.013	.027	.016	.829
D:						
.1	.588	.766	.608	.756	.868	1.880
1	.167	.285	.339	.507	.443	.733
20	.006	.012	.039	.075	.025	.460
E:						
.1	.557	.798	.896	.945	.888	.430
1	.187	.432	.701	.824	.603	.294
20	.007	.042	.149	.259	.081	.260
20:						
N:						
.1	.576	.578	.545	.705	.733	1.053
1	.184	.183	.144	.252	.309	1.188
20	.008	.007	.012	.023	.014	.901
U:						
.1	.592	.633	.069	.128	.776	20.470
1	.169	.196	.025	.050	.328	7.600
20	.007	.007	.008	.017	.015	1.203
M:						
.1	.579	.627	.249	.399	.773	5.170
1	.183	.192	.102	.186	.322	1.870
20	.008	.007	.017	.033	.015	.651
D:						
.1	.583	.634	.505	.671	.776	1.440
1	.186	.209	.290	.450	.346	.595
20	.007	.009	.038	.072	.018	.340
E:						
.1	.575	.711	.955	.977	.831	.106
1	.194	.371	.883	.938	.541	.072
20	.008	.041	.395	.566	.079	.069

Note: Median of 40 computer simulations is given for each case. The variable F_{st} is given for neutral electrophoretic markers and QTLs. Quantitative differentiation is given by Q_{st} and σ_b^2/σ_t^2 and compared with the value of σ_b^2/σ_t^2 predicted from QTL allelic differentiation. The proportion of σ_b^2 attributable to allelic differentiation at QTLs is also given ($\sum\sigma_{(i)}^2/\sigma_b^2$). Values >1 indicate that traits are less differentiated than predicted from single locus differentiation due to negative covariances among loci. N: neutral trait; U: uniform selection; M: moderate diversifying selection; D: diversifying selection; and E: extreme diversifying selection.

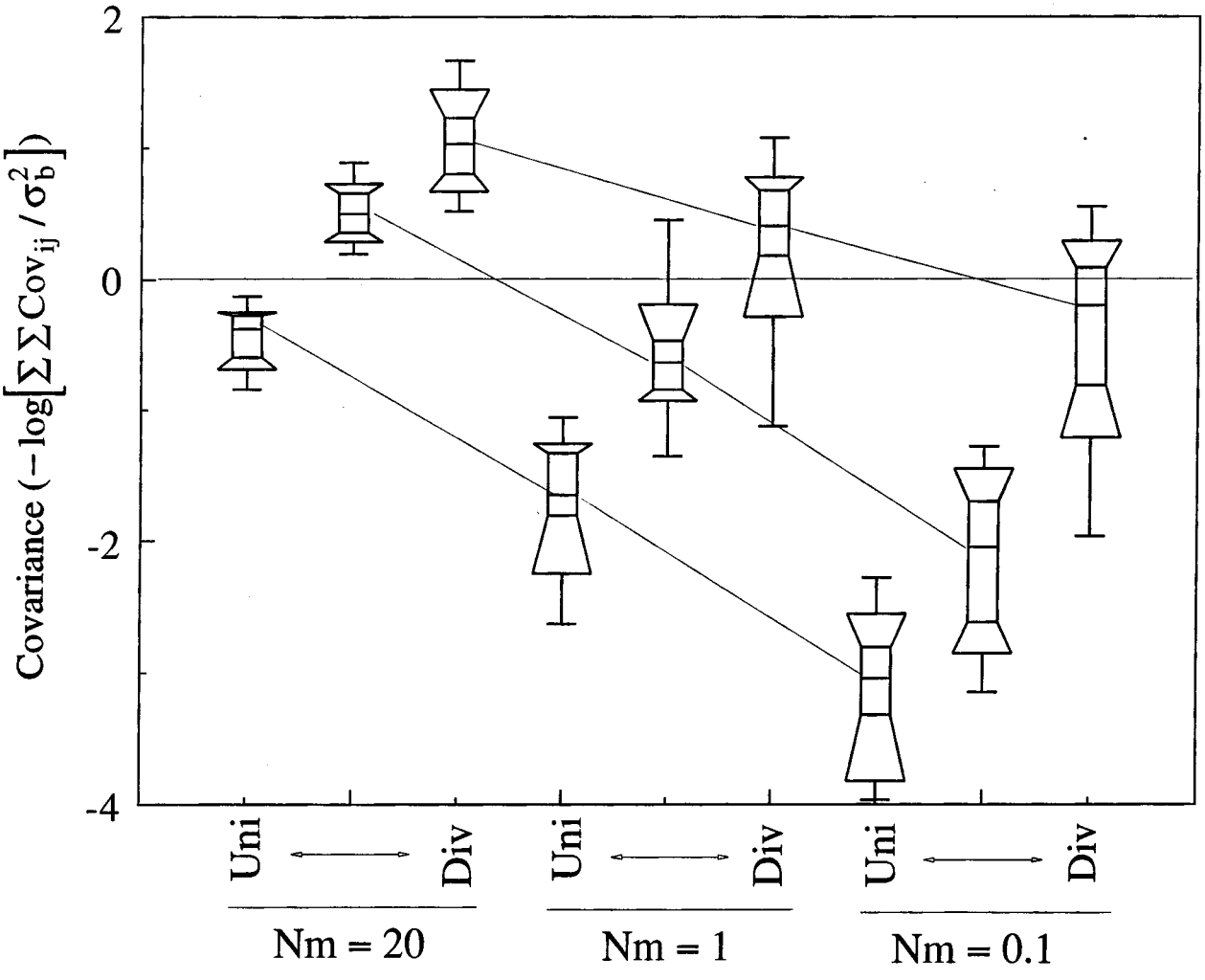

Figure 2: Box plots showing the contribution of covariances among QTL allelic frequencies to differentiation of trait values under various selection and migration regimes. For each combination of migration and selection regime, the distribution of $-\log[\sum\sum cov_{ij}/\sigma_b^2]$ over 40 simulated data sets is shown. Results are for the case of 20 QTLs. *Uni:* unifying selection (low $\sigma_{z^*}^2$); *Div:* diversifying selection (high $\sigma_{z^*}^2$).

in trait optima was greater than the neutral expected variance, positive correlations developed (fig. 2). Conversely, negative covariances were observed when the variance of trait optima was less than the neutral expectation. For a given migration rate (constant $2F_{st}\ \sigma_0^2$) increasing the variance in trait optima produced more positive covariances. By contrast, for a given selection regime (constant $\sigma_{z^*}^2$), restricting gene flow produced covariances that were more negative.

Correlations are particularly tight when few loci influence the trait, but the weaker correlations among loci with minor effects on the trait are offset by the disproportionately greater number of off-diagonal elements in $\mathbf{C}_b$. Thus, the trait generally showed a greater influence of covariances when many loci affected the trait than when few loci did (table 1). However, the percentage of σ_b^2 attributable to covariances of allelic frequencies was generally less than its theoretical maximum of $1 - 1/n$. Only in the case of extreme diversifying selection was the maximum approached. Even in these extreme cases, F_{st} of QTLs remained near that of neutral loci.

Surprisingly, however, when the variance in optima was equal to the neutral expectations, negative covariances developed (table 1), even though equation (6) predicted that no correlation would develop. In addition, F_{st} of QTLs was greater than for neutral loci in this case, in contrast to the expectation. Reasons for this discrepancy are discussed below.

Discussion

Differentiation of Loci and Traits

This model provides a unified framework in which to understand the relationship between differentiation of quantitative traits, and that of the underlying QTLs, for both unifying and diversifying selection. As selection acts on quantitative trait variation, it results not in changes of allelic differentiation at the underlying loci but, rather, creates covariances (linkage disequilibrium) across loci. This essentially decouples the level of differentiation of the trait from that of the QTLs.

The level of differentiation observed at single-gene markers, QTLs, and quantitative traits is expected to be equal only where such loci and traits are selectively neutral (Wright 1951, 1952; Rogers and Harpending 1983; Lande 1992; table 1). This occurs because, under neutrality, the overall covariances of allelic frequencies between populations are expected to be zero (table 1). Thus, equation (2) can be used to test the hypothesis that both electrophoretic markers and quantitative traits are neutral. In a series of studies covering diverse taxa, quantitative traits and single-gene electrophoretic markers show discordant levels of differentiation (Prout and Barker 1993; Spitze 1993; Podolsky and Holtsford 1995; Bonnin et al. 1996; Yang et al. 1996). Thus the hypothesis of neutrality must be rejected for either the electrophoretic markers or the quantitative traits. Assuming the electrophoretic markers to be neutral, this finding suggests that differentiation in the traits is caused by adaptation to the local environmental conditions experienced by each population.

However, given that the traits are locally adapted, the discordance between quantitative traits and single loci does not imply that allelic frequencies at QTLs are notably more or less divergent than at neutral markers. Thus equation (2) does not predict F_{st} of QTLs for selected traits as has sometimes been assumed. Local selection on quantitative traits acts not so much to alter the variance in allelic frequencies across populations (i.e., allelic differentiation) as to create covariances of allelic frequencies (linkage disequilibrium among populations). For example, common garden studies of conifers generally reveal levels of differentiation of $\sigma_b^2/\sigma_t^2 = 0.2$–$0.4$ (e.g., Campbell 1979; Yang et al. 1996). This corresponds to the case of extreme diversifying selection modeled here (table 1). By contrast, individual allozyme loci are minimally differentiated among conifer populations, indicating extensive gene flow (Hamrick and Godt 1990). Under these circumstances, allelic frequencies of QTLs need not be appreciably more differentiated than for neutral markers (F_{st} of about 0.05–0.1, table 1). Instead, covariances of allelic frequencies across loci can account for a high proportion of the between-population variance of the trait, such that QTLs will exhibit a pattern of variation that is similar to that at neutral loci.

The role of the covariance term in equation (3) helps to unify several previously reported findings regarding allelic variation for quantitative traits. For example, Lewontin (1984) created hypothetical examples of differentiation at QTLs to illustrate that differentiation at quantitative traits is usually easier to detect than that at individual loci. The discrepancy he points out can be accounted for by the high covariances of allelic frequencies across populations in his examples, which is formally equivalent to the variance in differentiation he discusses. However, quantitative traits and single loci differ not merely in the power to detect differentiation but also in the actual partitioning of variation within and among populations (table 1). Goldstein and Holsinger (1992) demonstrated that isolation by distance among subpopulations can permit allelic variation to be maintained in the face of strong stabilizing selection because strong negative covariances of allelic frequency occurs among isolated patches. In both cases, the discrepancy between trait and single locus differentiation can be described in terms of the between-population component of linkage disequilibrium (Ohta 1982).

The computer simulation presented in table 1 assumes linkage equilibrium within populations. However, stabilizing selection will act to create negative covariances of allelic frequencies within populations (Lande 1976). This linkage disequilibrium within populations does not negate the main result regarding the between-population component of linkage disequilibrium. Strong differentiation of quantitative traits is possible with only modest differentiation of QTL allelic frequencies, and vice versa (table 1). However, under strong selection within populations, the negative covariance of allelic effects will decrease the within-population variance (Goldstein and Holsinger 1992) and, thus, tend to increase the value σ_b^2/σ_t^2 that is observed for a given level of allelic frequency differentiation at QTLs.

Strong stabilizing selection within populations will also tend to fix the population for a multilocus genotype that closely matches the locally optimal phenotype. This will have two effects. First, it will decrease allelic variation within populations for QTLs, with a resulting increase in F_{st} (table 1). Second, since different multilocus combinations may give the same trait value, populations with similar optima may be fixed for different combinations of alleles (i.e., different multilocus genotypes), giving more negative covariances of allelic frequencies among populations. Thus slight negative covariances are some-

times observed where none were predicted by equation (6) (table 1, fig. 2).

There have been numerous instances of loci that exhibit unusually homogeneous or unusually differentiated gene frequencies across populations (e.g., Avise et al. 1979; Oakshott et al. 1982; Karl and Avise 1992). It has been suggested that such loci are unusual because they are experiencing natural selection (e.g., Cavalli-Sforza 1966; Lewontin and Krakauer 1973). While the interpretation of these loci with respect to selection remains unclear, we would not expect a discordant pattern of allelic frequency differentiation if such a locus were one of many additive genes contributing to (or linked to a locus contributing to) a quantitative trait under local selection. If these discordant loci are experiencing selection, then it seems likely that selection must act directly on those genes and not on a linked QTL.

Multilocus Measures of Population Differentiation

A number of authors have proposed that multilocus analyses of population structure at electrophoretic loci may provide an alternative way to relate electrophoretic surveys to ecologically relevant traits. For example, Lagercrantz and Ryman (1990) directly compared multilocus allozyme structure with quantitative trait differentiation in Norway spruce, *Picea abies,* and concluded that similar evolutionary forces have acted on allozymes and quantitative traits. Westfall and Conkle (1992) propose that multilocus analyses can improve discrimination of breeding zones for conifers in California. Such multilocus analyses can be performed in a variety of ways (Westfall and Conkle 1992; Yang and Yeh 1993; Kremer et al. 1997), most of which involve computing some estimator of the between-population variance-covariance matrix for the loci surveyed and taking eigenvectors of that matrix. The multilocus differentiation is then calculated as

$$F_{st(m)} = \boldsymbol{v}^{\mathrm{T}}\mathbf{C}_{\mathrm{b}}\boldsymbol{v}, \tag{7}$$

where $\boldsymbol{v}$ is the first eigenvector of the covariance matrix $\mathbf{C}_{\mathrm{b}}$ (Yang and Yeh 1993). Yeh et al. (1985) use such a method to show that a greater proportion of the total allozyme variance in lodgepole pine can be accounted for by differences among populations when this multilocus structure is accounted for than by analyses of single loci individually.

However, while multilocus analyses can improve discrimination among populations (Westfall and Conkle 1992), relating these analyses to quantitative traits that underlie local adaptation can be misleading. Though equations (4) and (7) are similar, the vector of allelic effects ($\mathbf{A}$ in eq. [4]) is determined a priori by the developmental mechanism that translates genotype into phenotype. By contrast, the eigenvectors of the matrix $\mathbf{C}_{\mathrm{b}}$ are determined a posteriori by the investigator, using a variance maximizing rotation that attributes as much variance as possible to differences among populations (Tatsuoka 1977). I applied such a multilocus measure of differentiation to simulated QTL allelic frequencies for a neutral trait (see neutral case, table 1). The multilocus F_{st} values consistently overestimate $\sigma_{\mathrm{b}}^2/\sigma_{\mathrm{t}}^2$ by up to fourfold (table 2). Therefore, the observation that similar proportions of variation in quantitative traits and multilocus discriminant scores are attributable to differences between populations is spurious.

Recently Kremer et al. (1997) proposed using the mean (rather than maximum) eigenvector of the between-population covariance matrix as an alternate measure of multilocus differentiation, which still takes covariances of allele frequencies across populations into account. Such a method could potentially provide a less-biased predictor of $\sigma_{\mathrm{b}}^2/\sigma_{\mathrm{t}}^2$ from allele frequency data because it would include all the dimensions of $\mathbf{C}_{\mathrm{b}}$, rather than just that dimension along which populations are maximally differentiated. However, such an estimator

Table 2: Comparison of population differentiation at quantitative traits and multilocus discriminant function of QTLs

		Observed			
		Five loci		*20 loci*	
Nm	*Expected* σ_b^2/σ_t^2	σ_b^2/σ_t^2	$F_{st(m)}$	σ_b^2/σ_t^2	$F_{st(m)}$
.1	.833	.787 (.532–.877)	.902 (.844–.949)	.754 (.579–.882)	.964 (.944–.990)
1.0	.333	.321 (.212–.465)	.520 (.392–.666)	.338 (.169–.479)	.721 (.621–.823)
20	.024	.039 (.012–.079)	.093 (.062–.149)	.047 (.012–.096)	.197 (.152–.269)

Note: For each combination of gene flow (Nm) and number of loci, the proportion of variance attributable to population differentiation is given for the trait, $\sigma_{\mathrm{b}}^2/\sigma_{\mathrm{t}}^2$, and the first discriminant function of allelic frequencies, $F_{st(m)}$ (median of 40 simulated data sets). Numbers in parentheses are ranges.

would have to be applied directly to the allele frequencies of the QTLs themselves. Unfortunately, these are usually not available, and there is no method available that could predict the covariance structure of QTL allelic frequencies from electrophoretic data (Strauss et al. 1992). Thus, electrophoretic surveys will provide only poor predictors of adaptive differentiation in quantitative traits.

The utility of electrophoretic studies will come instead from their ability to assess patterns of migration and population size and to estimate the action of migration/drift equilibrium, since differentiation of allelic frequencies is strongly influenced by these factors for both neutral markers and QTLs (table 1). The F_{st} values of neutral loci and QTLs are tightly correlated across a wide range of simulated scenarios (table 1). Where migration rates are low, allelic variation will be distributed primarily between populations for both neutral loci and QTLs. By contrast, with high migration rates, considerable allelic diversity will be harbored within each population at both neutral and QTL loci. Therefore, while common garden studies will provide the best practical measures of quantitative trait differentiation, electrophoretic surveys will likely provide a more accurate picture of the distribution of allelic variation underlying those traits.

Acknowledgments

I am very grateful to J. B. Mitton, R. H. Podolsky, J. Rehfeldt, K. Ritland, K. Spitze, and F. C. Yeh for discussions and comments on an earlier draft of this article. This work was supported by a National Science Foundation Doctoral Dissertation Improvement Grant.

Literature Cited

Avise, J. C., M. H. Smith, and R. K. Selander. 1979. Biochemical polymorphism and systematics in the genus *Peromyscus.* VII. Geographic differentiation in members of the *trueii* and *maniculatus* species groups. Journal of Mammalogy 60:177–192.

Bonnin, I., J. M. Prosperi, and I. Olivieri. 1996. Genetic markers and quantitative genetic variation in *Medicago truncatula* (Leguminosae): a comparative analysis of population structure. Genetics 143:1795–1805.

Bradshaw, A. D. 1984. Ecological significance of genetic variation between populations. Pages 213–228 *in* R. Dirzo and J. Sarukhan, eds. Perspectives on plant population ecology. Sinauer, Sunderland, Mass.

Campbell, R. K. 1979. Genecology of Douglas-fir in a watershed in the Oregon Cascades. Ecology 60:1036–1050.

Cavalli-Sforza, L. 1966. Population structure and human evolution. Proceedings of the Royal Society of London B, Biological Sciences 164:362–379.

Falconer, D. S. 1989. Introduction to quantitative genetics. 3d ed. Longman, New York.

Goldstein, D. B., and K. E. Holsinger. 1992. Maintenance of polygenic variation in spatially structured populations: roles for local mating and genetic redundancy. Evolution 46:412–429.

Hamrick, J. L., and M. J. Godt. 1990. Allozyme diversity in plant species. Pages 43–63 *in* A. H. D. Brown, M. T. Clegg, A. L. Kahler, and B. S. Weir, eds. Plant population genetics, breeding and genetic resources. Sinauer, Sunderland, Mass.

Jordan, N. 1991. Multivariate analysis of selection in experimental populations derived from hybridization of two ecotypes of the annual plant *Diodia teres* W. (Rubiaceae). Evolution 45:1760–1772.

Karl, S. A., and J. C. Avise. 1992. Balancing selection at allozyme loci in oysters: implications from nuclear RFLP's. Science (Washington, D.C.) 256:100–102.

Kremer, A., A. Zanetto, and L. Ducousso. 1997. Multilocus and multitrait measures of differentiation for gene markers and phenotypic traits. Genetics 145:1229–1241.

Lagercrantz, U., and N. Ryman. 1990. Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozymic variation. Evolution 44:38–53.

Lande, R. 1976. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genetical Research 26:221–235.

———. 1992. Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. Evolution 46:381–389.

Lascoux, M., J. Thorsen, and U. Gullberg. 1996. Population structure of a riparian willow species, *Salix viminalis* L. Genetical Research 68:45–54.

Lewontin, R. C. 1984. Detecting population differences in quantitative characters as opposed to gene frequencies. American Naturalist 123:115–124.

Lewontin, R. C., and J. Krakauer. 1973. Distribution of gene frequency as a test of the selective neutrality of polymorphisms. Genetics 54:595–609.

Oakshott, J. G., J. B. Gibson, P. R. Anderson, W. R. Knibb, D. G. Anderson, and G. K. Chambers. 1982. Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. Evolution 36:86–96.

Ohta, T. 1982. Linkage disequilibrium due to random genetic drift in finite subdivided populations. Proceedings of the National Academy of Sciences of the USA 79:1940–1944.

Paterson, A. H., S. Damon, J. D. Hewitt, D. Zamir, H. D. Rabinowitz, W. E. Lincoln, E. S. Lander, and S. D. Tanksley. 1991. Mendelian factors underlying quanti-

tative traits in tomato: comparison across species, generations and environments. Genetics 127:181–197.

Podolsky, R. H., and T. P. Holtsford. 1995. Population structure of morphological traits in *Clarkia dudleyana.* I. Comparison of F_{st} between allozymes and morphological traits. Genetics 140:733–744.

Prout, T., and J. S. F. Barker. 1993. *F* statistics in *Drosophila buzzatii:* selection, population size and inbreeding. Genetics 134:369–375.

Rogers, A. R., and H. C. Harpending. 1983. Population structure and quantitative characters. Genetics 105: 985–1002.

Schemske, D. W. 1984. Population structure and local selection in *Impatiens pallida* (Balsaminaceae), a selfing annual. Evolution 38:817–832.

Spitze, K. 1993. Population structure in *Daphnia obtusa:* quantitative genetic and allozyme variation. Genetics 135:367–374.

Strauss, S. H., R. Lande, and G. Namkoong. 1992. Limitations of molecular-marker aided selection in forest tree breeding. Canadian Journal of Forest Research 22: 1050–1061.

Tatsuoka, M. M. 1977. Multivariate analysis: techniques for educational and psychological research. Wiley, New York.

Turelli, M. 1984. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theoretical Population Biology 25:138–193.

Westfall, R. D., and M. T. Conkle. 1992. Allozyme markers in breeding zone designation. New Forests 6: 279–309.

Wright, S. 1951. The genetical structure of populations. Annals of Eugenics 15:323–354.

———. 1952. The theoretical variance within and among subdivisions of a population that is in a steady state. Genetics 37:312–321.

Yang, R. C., and F. C. Yeh. 1993. Multilocus structure in *Pinus contorta* Dougl. Theoretical and Applied Genetics 87:568–576.

Yang, R. C., F. C. Yeh, and A. D. Yanchuk. 1996. A comparison of isozyme and quantitative genetic variation in *Pinus contorta* ssp. latifolia by F_{st}. Genetics 142: 1045–1052.

Yeh, F. C., W. M. Cheliak, B. P. Dancik, K. Illingworth, D. C. Trust, and B. A. Pryhitka. 1985. Population differentiation in lodgepole pine *Pinus contorta* ssp. *latifolia:* a discriminant analysis of allozyme variation. Canadian Journal of Genetics and Cytology 27:210–218.

Associate Editor: Kent E. Holsinger