Inference From Clines Stabilized by Frequency-Dependent Selection

James Mallet* and Nick Barton†

**Department of Entomology, Mississippi State University, Mississippi State, Mississippi 39762, and †Galton Laboratory, Department of Genetics and Biometry, University College, London NW1 2HE, England*

Manuscript received November 7, 1988
Accepted for publication April 28, 1989

ABSTRACT

Frequency-dependent selection against rare forms can maintain clines. For weak selection, s, in simple linear models of frequency-dependence, single locus clines are stabilized with a maximum slope of between $\sqrt{s}/\sqrt{8}\ \sigma$ and $\sqrt{s}/\sqrt{12}\ \delta$, where σ is the dispersal distance. These clines are similar to those maintained by heterozygote disadvantage. Using computer simulations, the weak-selection analytical results are extended to higher selection pressures with up to three unlinked genes. Graphs are used to display the effect of selection, migration, dominance, and number of loci on cline widths, speeds of cline movements, two-way gametic correlations ("linkage disequilibria"), and heterozygote deficits. The effects of changing the order of reproduction, migration, and selection, are also briefly explored. Epistasis can also maintain tension zones. We show that epistatic selection is similar in its effects to frequency-dependent selection, except that the disequilibria produced in the zone will be higher for a given level of selection. If selection consists of a mixture of frequency-dependence and epistasis, as is likely in nature, the error made in estimating selection is usually less than twofold. From the graphs, selection and migration can be estimated using knowledge of the dominance and number of genes, of gene frequences and of gametic correlations from a hybrid zone.

GENE frequency clines can be stabilized by various forms of natural selection. Theoreticians have mainly investigated cases where geographic variation in selection pressures (HALDANE 1948; FISHER 1950; SLATKIN 1973, 1975; MAY, ENDLER and MCMURTRIE 1975; ENDLER 1977) or heterozygote disadvantage (BAZYKIN 1969; BARTON 1979) maintain clines. Clines can also be stabilized purely by epistasis (BAZYKIN 1973), and by frequency-dependent selection if the rare form is at a disadvantage (MALLET 1986a). Heterozygote disadvantage, epistasis, and frequency-dependent selection can all maintain "tension zones" (HEWITT 1988). Tension zones differ from ecologically determined clines in that they are not rooted to environmental gradients, but tend to move until trapped by dispersal barriers or areas of low population density (BAZYKIN 1969; BARTON 1979).

Tension zones connect groups of populations at different stable genetic equilibria (often adaptive peaks); thus they form an important component of WRIGHT's (1932) "shifting balance" model for the evolution of new stable equilibria. Of course, natural tension zones can form by secondary contact between already-diverged populations; but if divergence by the shifting balance does occur in parapatry, a tension zone will be present between the nascent equilibrium and the equilibrium from which it evolved. The shifting balance can be divided into four phases. (1) Gene frequencies drift against the force of selection in a local group of demes [or, in a continuous model, a local group of "neighborhoods" (WRIGHT 1969)] linked by migration. (2) Once a critical, unstable equilibrium has been passed because of drift, natural selection continues the process of pushing these local demes (or neighborhoods) towards a new stable equilibrium. (*3*) Provided a critical number of demes (or neighborhoods) have reached the new equilibrium, this equilibrium will be protected from the old form by a set of clines which constitute a stable, though mobile, tension zone around the new forms. (4) Such zones can move out to spread the new equilibrium or collapse inwards to cause the loss of the new form, depending on the fitness of the new form, as well as on variation in population density and dispersal (BARTON 1979; ROUHANI and BARTON 1987). This fourth process of selection across a tension zone has the same effect as Wright's "interdemic selection." In a demic model with low migration between demes, the entire tension zone is essentially between a pair of adjacent demes (SLATKIN 1981; WALSH 1982; LANDE 1985). In nature, more or less continuous tension zones spanning many demes are common (BARTON and HEWITT 1985; HEWITT 1988); the shifting balance under these conditions has been modeled by ROUHANI and BARTON (1987). The potential for rapid, deterministic cline movement to preserve and spread new equilibria has not generally been appreciated. This

The publication costs of this article were partly defrayed by the payment of page charges. This article must therefore be hereby marked "*advertisment*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

kind of selection across continuous tension zones, as well as interdemic selection, makes the shifting balance much more likely than is often assumed, in the same way that individual selection allows the spread of rare advantageous mutations within a population (WRIGHT 1980). The study of tension zones is therefore important for understanding genetic divergence, including chromosomal evolution (BAZYKIN 1959) and some kinds of speciation (WRIGHT 1982; BARTON and CHARLESWORTH 1984).

Warning and mimetic colors of insects often differ greatly between regions. Stable clines (tension zones) in warning pattern can be maintained by density-dependent predator attacks: when a form is rare, it is selected against because predators do not recognize it as being unpalatable. This predator behavior gives rise to frequency-dependent selection which causes Müllerian mimicry between unpalatable species, as well as lack of color polymorphism within unpalatable species (MALLET and SINGER 1987; ENDLER 1988). There have been no previous theoretical analyses of clines maintained by frequency-dependent selection, although such theory would be useful for analyzing many hybrid zones between mimetic and warningly colored insect taxa: there are thousands of such hybrid zones known in neotropical butterflies alone (*e.g.*, BROWN, SHEPPARD and TURNER 1974; BROWN 1979); similar clines are found in bumblebees (OWEN and PLOWRIGHT 1980; P. H. WILLIAMS personal communication) and soldier beetles (MCLAIN 1988) among many other insects. Of course, other mechanisms of selection could also be involved: however, the presence of tight Müllerian mimicry between many of the species is good evidence that frequency-dependence is operating within these species. Frequency-dependent selection is likely to stabilize a variety of other clines: for example fronts between strains of bacteria with and without the ability to produce allelochemicals (LEVIN 1988), and clines between races having assortative mating (MOORE 1979) or sexual selection provided that the rare form is at a disadvantage (LANDE 1982; MOORE 1987; JOHNSON, MURRAY and CLARKE 1987). There is also a close relationship, demonstrated below, between clines maintained by frequency-dependent selection and those maintained, like many chromosomal clines, by heterozygous disadvantage.

We here analyze tractable cases of clines at one locus with weak frequency-dependent selection, and use computer simulations to investigate more complex cases of strong selection on one, two and three loci. The immediate purpose of this work is to provide models for analyzing hybrid zones in Heliconius (Lepidoptera): these models enable one to infer both selection and gene flow without recourse to field experiments [see BARTON (1982) and SZYMURA and BARTON (1986) for examples of the method applied to other clines].

CLINES AT A SINGLE LOCUS UNDER WEAK SELECTION

When selection is weak, gene flow may be approximated by diffusion at a rate σ^2, so that over a linear dimension x, the gene frequency p of a locus A changes with time t as follows (HALDANE 1948; NAGYLAKI 1975):

$$\frac{dp}{dt} \approx \frac{\sigma^2}{2} \frac{\delta^2 p}{\delta x^2} + S(p) \tag{1}$$

where σ^2 is the variance in parent-offspring distance along some axis, and $S(p)$ is the effect of selection. A simple case of a diploid frequency-dependent $S(p)$ has intermediate heterozygote fitness:

$$\left.\begin{aligned} W_{AA} &= 1 - 2s\left(f_{aa} + \frac{1}{2} f_{Aa}\right) \\ W_{aa} &= 1 - 2s\left(f_{AA} + \frac{1}{2} f_{Aa}\right) \\ W_{Aa} &= (w_{aa} + w_{AA})/2 = 1 - s \end{aligned}\right\} \tag{2}$$

where W is fitness, f is the frequency of genotypes, and $0 \leq s \leq 0.5$ is a constant determining the strength of frequency-dependence. The selection coefficient of $0 \leq 2s \leq 1$ was chosen to make the similarity between frequency-dependent selection and heterozygous disadvantage self-evident, as explained below. At low selection pressures, mean fitness ≈ 1 and these fitnesses lead to:

$$S(p) \approx spq(p - q). \tag{3}$$

This is identical to $S(p)$ for a heterozygous disadvantage of s (BAZYKIN 1969). Thus the homozygote fitnesses in the frequency-dependent model play little role in the dynamics under the diffusion approximation, which applies when selection is weak. The differences between heterozygous disadvantage and frequency-dependent selection enter only through the mean fitness, which will be close to 1 for small s. Equation 1 can then be integrated to give, at equilibrium:

$$p \approx \frac{1}{2}\left\{1 + \tanh\left[\frac{(x - x_0)\sqrt{(2s)}}{2\sigma}\right]\right\}$$

where x_0 is the point at which $p = 0.5$ (BAZYKIN 1969; BARTON 1979). Figure 1A shows this integral. The maximum gradient for this frequency-dependent model ($dp/dx \approx \sqrt{s}/\sqrt{8}\,\sigma$) is of course again the same as that for heterozygous disadvantage. Heterozygous disadvantage can be thought of as leading to frequency-dependent selection at the haploid level: the rarer gene encounters the opposite allele more fre-

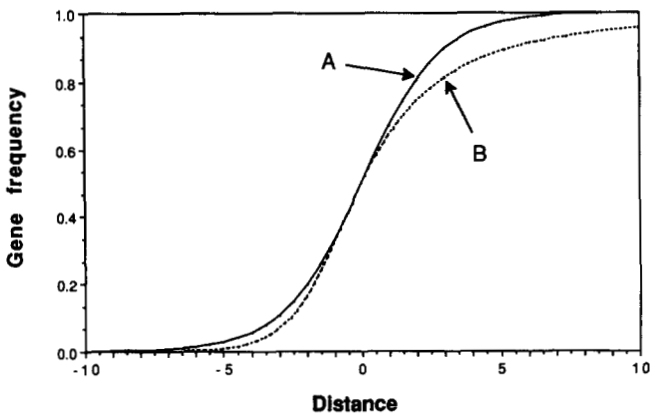

FIGURE 1.—Single locus clines at equilibrium under weak selection. Curve *A*, a codominant locus, Equation 3. Equation 6 is similar: selection is approximately as effective in the center, but only 1/2 as effective at the edge: the cline will be $\sqrt{(2/3)}$ as steep. Curve *B*, a dominant locus, equation (8). Distance is measured in standardized units of $\sigma/4\sqrt{s}$.

quently, and so has a lower average fitness.

The model in (2) is unrealistic because the fitness of heterozygotes is expected to depend on the frequencies of similar and different phenotypes, rather than being a simple average of the homozygote fitnesses. The possibility of dominance also needs to be incorporated. With warning color, heterozygotes can be remembered as a separate phenotype, although "generalization" between heterozygotes and homozygotes (the tendency for predators to confuse similar phenotypes) is likely. Predator generalization between all three genotypes may also occur, but will not cause differential selection, and so is ignored. We use a linear frequency-dependent model for three reasons: a linear model can be viewed as a regression through the actual frequency-dependent function; the actual nonlinear selection function is not known; and linear and nonlinear functions produce similar clines (MALLET 1986a). Selection is assumed not to affect population density, so that only genotype frequencies need to be modeled. This is likely for a butterfly with strong density-dependent factors controlling population size in the larval stage, as is apparently true in Heliconius (GILBERT and SMILEY 1978; MALLET 1984). In the simplest case, selection reduces the frequency of a genotype in proportion to its phenotypic rarity $(1 - f)$, where f is the frequency of that genotype and of similar phenotypes. For example, $W_{AA} = 1 - 2s[1 - f_{AA} - h f_{Aa}] = 1 - 2s[f_{aa} + (1 - h)f_{Aa}]$, where the dominance of A (equivalent to predator generalization between Aa and AA) is given by $0 \leq h \leq 1$. The complete model is then:

$$\left.\begin{aligned} W_{AA} &= 1 - 2s_{AA}\,[f_{aa} + (1 - h)\,f_{Aa}] \\ W_{aa} &= 1 - 2s_{aa}\,(f_{AA} + h\,f_{Aa}) \\ W_{Aa} &= 1 - 2s_{Aa}\,[h\,f_{aa} + (1 - h)\,f_{AA}] \end{aligned}\right\} \quad (4)$$

where the individual selection coefficients, $2s$, in warning color depend on the memorability of that pattern's genotype. The three selection coefficients represent the actual reduction in fitness due to rarity for each genotype. If $s_{Aa} > s_{AA}$ and $s_{Aa} > s_{aa}$, heterozygotes have a disadvantage—this may occur with warning color if the hybrid phenotype is less memorable. In warning color, if $s_{AA} < s_{aa}$ the A pattern is more memorable than a, and interdemic selection between two color patterns might be expected to move clines so as to fix A over the whole range [see BARTON (1979) for similar movement of clines with heterozygous disadvantage and selection against one homozygote]. The simplest case is to assume that there are no differences in memorability of the phenotypes ($s_{AA} = s_{Aa} = s_{aa} = s$); then

$$S(p) \approx 2spq[(1 - h)p(2p^2 - 1) + hq(1 - 2q^2)]. \quad (5)$$

If the locus A is codominant, $h = 1/2$, and

$$S(p) \approx spq(p^2 + q^2)(p - q). \quad (6)$$

Again $S(p)$ approximates that for heterozygous disadvantage, differing only by a factor of $(p^2 + q^2)$, which always lies between 1 (at the edge of the cline) and 1/2 (in the center). The cline shape therefore differs little from that in (3) (Figure 1A), except that the maximum gradient obtained by integration ($dp/dx \approx \sqrt{s}/\sqrt{12}\sigma$) is $\sqrt{(2/3)}$ as great. This is confirmed in the simulations below for low s. In contrast, dominance ($h = 1$) gives:

$$S(p) \approx 2spq^2(1 - 2q^2). \quad (7)$$

A cline set up with this selection regime will move because of the interdemic selective asymmetry due to dominance (MALLET 1986a). It is hard to find the shape of a moving cline because the velocity cannot, in general, be found analytically. We can find the approximate cline shape by supposing that there is a selective asymmetry which just suffices to prevent movement. In nature also, movement may often be prevented by a dispersal or selection gradient which will produce similar effects. The necessary asymmetry is given by:

$$S(p) \approx 2spq^2\,(4/5 - 2q^2). \quad (8)$$

Setting $dp/dt = 0$ in (1) gives for this model:

$$x' - x'_0 = \sqrt{(15/2)} \int_{0.5}^{q} \frac{dq}{q(1 - q)\sqrt{[q(4 + 5q)]}}$$

where distance is now measured in standardized units of $x' = \sigma/2\sqrt{s}$. This can be integrated numerically to give the cline in Figure 1B. The cline is asymmetrical: on the side where dominants are common, most recessive alleles will be found in heterozygotes, and so will escape selection, giving a long tail; on the recessive

side of the cline on the other hand, selection weeds out the rare dominant alleles even when in heterozygotes, giving an abrupt tail. HALDANE (1948) showed that similar asymmetries are expected with dominance in clines responding to environmental discontinuities. Because there are only two phenotypes, we might expect the gradient in the center of the cline to be similar to that of (3): this is confirmed in the simulations.

CLINES WITH ARBITRARILY STRONG SELECTION; 1–3 LOCI

Methods: In the above results $S(p)$ was calculated assuming that $\bar{W} \approx 1$, and that genotypic frequencies are in Hardy-Weinberg equilibrium. Neither holds under strong selection. Analytic results also become difficult with more than one locus. We therefore performed computer simulations of clines on a linear array of demes to investigate how selection affects the maximum gradient, speed of movement, heterozygote deficiency within demes (F), and linkage disequilibria ($R = D/\sqrt{[p_A q_a p_B q_b]}$, where $D = f_{AB} - p_A p_B$ and f_{AB} is the frequency of AB gametes).

The simulations employ the exact linear frequency-dependent selection given by (4), with either complete dominance ($h = 1$, also used by MALLET [1986a]) or codominance ($h = 1/2$). We modeled a one-dimensional cline in a series of demes connected by migration. In MALLET (1986a), "stepping-stone" (KIMURA and WEISS 1964) migration between adjacent demes was used. However, this causes inaccuracies when $s/\sigma \approx 1$ or greater because the clines become grossly discontinuous. Here we use binomial migration, a generalization of the stepping-stone model, with the "number of trials" $= 2n$ demes. The proportion arriving in a deme i demes away from the source is:

$$\frac{(2n)!}{(n-i)!\,(n+i)!}\,\alpha^{(n-i)}\,(1-\alpha)^{(n+i)}.$$

For symmetric migration as used here, $\alpha = 1 - \alpha = 1/2$, and the variance of parent-offspring distances, σ^2, in simply n. This migration is approximately Gaussian for large n. Beyond each edge of the deme array there were assumed to be n imaginary demes with genotypic frequencies identical to those of the edge deme. These imaginary demes acted as sources for the immigrants that arrived within the deme array. We checked for approximate continuity by doubling n: if the resulting cline widths did not scale linearly with σ, we used a still higher level of migration. Edge effects were eliminated similarly; if the cline shape was affected by changing the total number of demes, the demes were increased until no further effect was noticed.

Strong selection on warning color genes is likely in the field [each gene may consist of many tightly linked loci: see MALLET (1986a, 1989) and MALLET and BARTON (1989)]. Therefore we ran simulations with selection pressures, $2s$, of 0.05, 0.1, 0.2, 0.3, 0.5, 0.75, 0.9 and 0.99. Migration levels, σ^2, used for each selection pressure were 10, 10, 15, 15, 30, 30, 30 and 30, respectively. The fitness of a genotype is assumed to be the product of the separate fitnesses at each locus. Such multiplicative fitnesses can be justified if selection occurs independently on each locus: the probability of survival is then equal to a product of the separate probabilities of escape due to each locus. Additive fitnesses could have been used, but there is very little difference between these modes with weak selection.

In reality, selection will of course be much more complex. Consider warning color: a number of parameters will characterize phenotypic selection, the frequency-dependence will be nonlinear, and epistasis, which is modeled below for some special cases, will produce strong interactions between the fitnesses at individual loci (*e.g.*, MALLET 1989). Inference from these simplified models will therefore give an "effective" selection pressure that is dependent on this idealized model situation.

For the most part, we assume that genotypes or phenotypes have equal frequency-dependence (in warning color, the phenotypes would then be equally memorable) so that a single per locus selection pressure, $2s$, characterizes all phenotypes within and between different loci. We repeated the simulations for one-locus clines, and for hybrid zones involving two or three loci, with various combinations of dominance for each locus ($h = 1/2$ or 1). The loci are assumed to be unlinked throughout. Measurement of p, F, R, and third order D [as defined by SLATKIN (1972) and BARTON (1982)] took place immediately before reproduction in each generation. The order of processes was normally assumed to be: reproduction (consisting of mating, offspring production, and growth to adulthood)-migration-selection. For Heliconius this seems approximately valid: adults eclose, immediately disperse, and then set up home ranges (MALLET 1986b,c), and many vertebrate species will be similar. With strong selection the order of processes matters [see also MANI (1985)] so we have also investigated another possibility: reproduction–selection–migration. For a few special cases we have also demonstrated the effect of removing disequilibria at meiosis, and of combining epistasis with frequency-dependent selection.

Epistasis, which is likely in many natural hybrid zones (*e.g.*, in Heliconius, see MALLET [1989]) could be modelled in a wide variety of ways, but the main feature needed here is an interaction between genes such that the fitness of pure phenotypes is greater. For simplicity, we suppose that the two pure warning

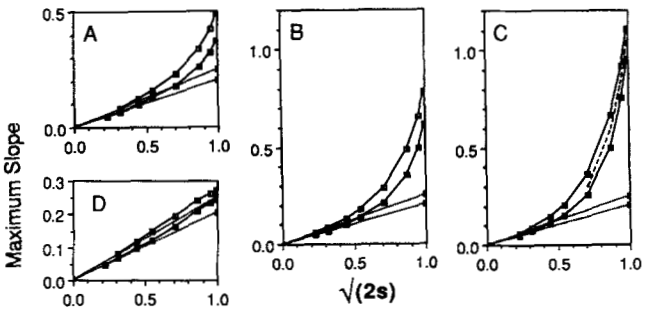

FIGURE 2.—Maximum slope (standardized by multiplying by σ) in simulated tension zones plotted against selection, $\sqrt{(2s)}$. (A) 1 locus; (B) 2 loci; (C, D) 3 loci; (A)–(C) all have the standard order reproduction–migration–selection; (D) has an alternative order of reproduction–selection–migration. The lines represent all dominant (open squares), and all codominant (filled squares) clines. In (C) the mean slope (dashed line) of clines in a zone with 1 codominant locus and 2 dominant loci is also plotted. The low s approximations for dominant genes ($\sqrt{[s/8]}$, upper dotted line), and codominant genes ($\sqrt{[s/12]}$, lower dotted line) are also shown in each figure.

color patterns are equally fit, but that any hybridization with the other pattern reduces fitness by an amount e for each whole phenotypic element changed, and $e/2$ for each heterozygous element at codominant loci. Suppose a mixture of one codominant locus, *A*, and two completely dominant loci, *B* and *C* interact in a tension zone (as occurs in Heliconius). Then *AA B– C–* and *aa bb cc* are both pure patterns having maximum fitness 1; *AA B– cc*, *AA bb C–*, *AA bb cc*, *aa B– C–*, *aa B– cc* and *aa bb C–* are one step away from pure, with fitness $1 - e$; *Aa B– C–* and *Aa bb cc* are half a step away, and have fitness $1 - e/2$; *Aa B– cc* and *AA bb C-* are 1½ steps away, and have fitness $(1 - e)(1 - e/2)$. Note that a result of this conceptually simple model of "epistasis" is an automatic heterozygote disadvantage of $e/2$ for codominant loci: their heterozygotes are always deviants from the pure pattern. Simulations have been run for a few cases using various mixtures of this epistatic model and the frequency-dependent model, with the two types of fitnesses combined multiplicatively.

Cline shape: We have seen that the maximum gradient of a single locus cline under weak selection is proportional to $\sqrt{s/\sigma}$. When selection is increased, the diffusion approximation breaks down and the cline steepens (Figure 2A) because second order terms in s become important. As expected from the arguments above, dominance steepens the clines because the diversity of phenotypes decreases.

An example of a simulated tension zone involving three loci is shown in Figure 3. The shape of each cline in a multilocus hybrid zone is also affected by the number of loci involved. As before, the maximum slope increases linearly with $\sqrt{(2s)}$ at low selection pressures (Figure 2, B and C). At high selection pressures, the approximation breaks down, and the slope increases even more steeply than for a single locus.

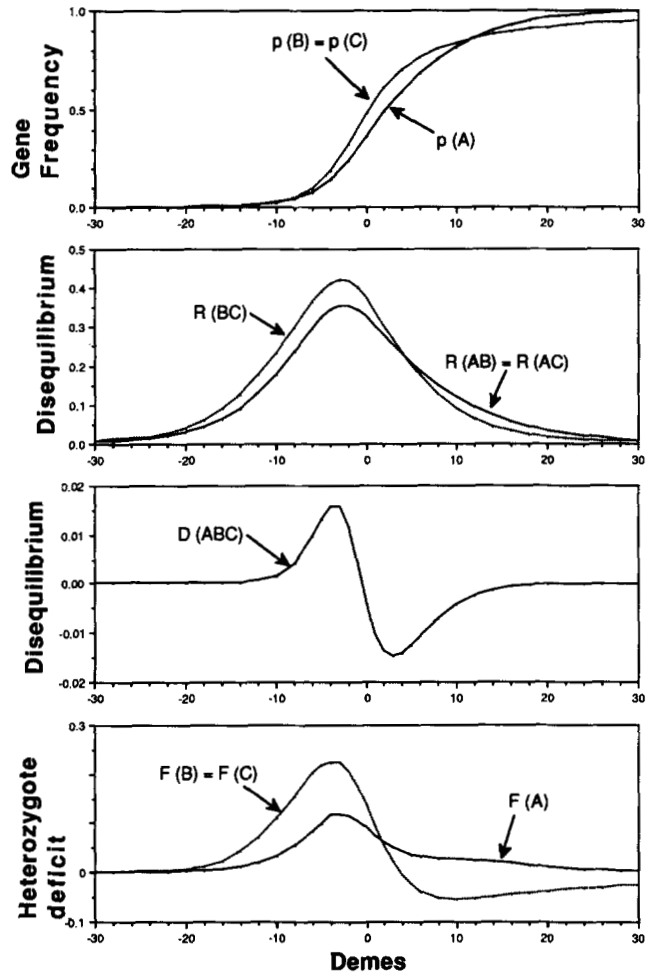

FIGURE 3.—Simulated three-locus tension zone consisting of one codominant locus (*A*) and two completely dominant loci (*B*, *C*). Such a hybrid zone occurs in Heliconius (MALLET and BARTON 1989). The clines are stabilized by frequency-dependent selection, $2s = 0.5$ per locus, and migration, $\sigma^2 = 15$ demes per generation. The figure shows variation along a demic continuum of all the parameters needed to reconstruct the genotypic distributions. Only the central section of the simulation is shown out of a total of 150 demes modeled. The simulation has been running for 250 generations since contact between fixed populations, has stabilized in shape and is moving to the left at a constant speed of ~0.20 demes (~0.051 dispersal distances, σ) per generation. At lower selection pressures such a zone would separate into a stationary *A* cline and a pair of rapidly moving clines, *B* and *C*. Top to bottom: gene frequencies; two-way disequilibria, R_{IJ}; three-way disequilibrium D_{ABC} (defined as in SLATKIN 1972); and heterozygote deficits, F_{IS}.

This is partly due to second order terms in the multiplicative fitnesses; in addition, gametic correlations ("linkage disequilibria") are generated both by the selection model, and by migration (see below). These disequilibria tend to steepen the cline still further because selection on one locus will act on associated loci (SLATKIN 1975; BARTON 1983). This explanation can be tested by artificially removing the disequilibria at meiosis in each simulated generation. The slopes of such clines are considerably reduced compared with those for the same loci in tension zones where disequilibria are allowed to accumulate naturally, though

TABLE 1

Effect of removing disequilibria at meiosis

$\sigma \times$ Max. (dp/dx)	Speed/σ	Max. R	Max. F	Min. F
Three loci, normal:				
A 0.580	0.053	*AB/AC* 0.604	*A* 0.315	*A* 0.000
B,C 0.597		*BC* 0.620	*B,C* 0.394	*B,C* −0.079
Mean 0.591		Mean 0.609	Mean 0.368	
Three loci, disequilibrium removed:				
A 0.410	0.059	*AB/AC* 0.278	*A* 0.190	*A* 0.000
B,C 0.488		*BC* 0.325	*B,C* 0.345	*B,C* −0.085
Mean 0.462		Mean 0.293	Mean 0.294	
One locus, codominant:				
0.260	0.000		0.075	0.000
One locus, dominant:				
0.337	0.083		0.283	−0.103

Results of simulations of three- and one-locus tension zones with $2s = 0.75$ and $\sigma^2 = 30$, using the standard order of processes: reproduction–migration–selection. In the three-locus results, *A* is codominant, *B* and *C* are completely dominant, as in Figure 3.

they still remain steeper than in the equivalent single locus clines (Table 1); the remainder of the increase in slope is due to the non-additive epistasis inherent in multiplicative fitnesses and to the disequilibria produced by migration within each generation. In the center of multilocus hybrid zones, disequilibria cause loci to be selected as though they were a linked block, leading to a sharp step in gene frequency; at the edges, each locus is selected on its own, giving long tails of introgression on either side of the central step. This will also be true at low s, provided there are many loci. Here, the effect is only appreciable at high s because the small number of unlinked loci are otherwise "weakly coupled" (see BARTON 1983).

As in Figure 1, dominant clines are asymmetric with long tails of recessive alleles on the side of the cline where dominants are most common (Figure 3). Due to disequilibria with dominant loci, codominant clines in mixed tension zones also become slightly asymmetric (Figure 3). Alone, they are completely symmetrical (Figure 1).

Cline movement: Clines maintained by heterozygous disadvantage tend to move with constant speed if one homozygote is fitter than another (BAZYKIN 1969; BARTON 1979); this applies also to the similar clines with frequency-dependent selection on codominant loci. In contrast, dominant clines move even in the absence of differential selection of phenotypes (Figures 3 and 4), because dominance introduces a selective asymmetry on the genotypes (MALLET 1986a). This cline movement due to dominance is little affected by the numbers of dominant loci; however, if one of the loci is codominant, the movement is considerably slowed. Clines differing in dominance tend to move apart below some critical s; at higher s, disequilibria and epistasis cause the clines to be attracted together (Figures 3 and 4) (SLATKIN 1975). For instance, where one codominant and two dominant loci interact in a tension zone, the critical $2s$ is between 0.4 and 0.5. We do not analyse here cases of a selective difference between homozygotes [for a treatment applicable to codominant frequency-dependence, see BARTON (1979)]; such movement like that due to dominance, is at a speed approximately proportional to $\sigma\sqrt{s}$ (Figure 4).

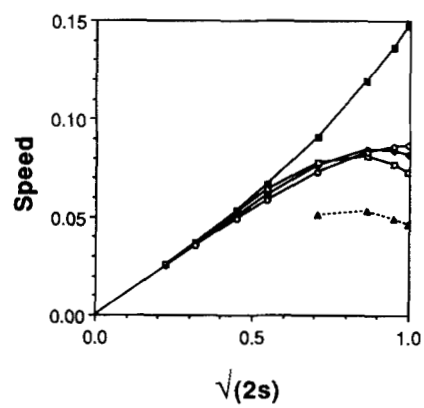

FIGURE 4.—Cline movement due to dominance. Curves show movement speeds of tension zones (measured in units of dispersal distances, σ, per generation) plotted against selection pressure, $\sqrt{(2s)}$. Open symbols and solid lines: circles, 1 dominant locus; diamonds, 2 dominant loci; squares, 3 dominant loci; triangles and dashed lines, 1 codominant and 2 dominant loci. All are for the standard order (reproduction–migration–selection) except for the upper curve (filled squares) which is for 3 dominant genes with the order reproduction–selection–migration. Cline movement due to advantage of one of the homozygotes also shows an approximately linear dependence on $\sigma\sqrt{s}$ (BARTON 1979).

Gametic correlations: Gene flow across a set of selected clines will produce correlations both within loci–heterozygote deficiencies, F–and between loci–linkage, or gametic, disequilibria, measured by the correlation coefficient R (BARTON 1982, 1983; SZYMURA and BARTON 1986). If gene flow is kept constant and selection is increased, the cline steepens, and migration moves genotypes across a sharper gradient, causing an increase in disequilibrium. In contrast, increasing gene flow while keeping selection constant rescales the cline width without altering R.

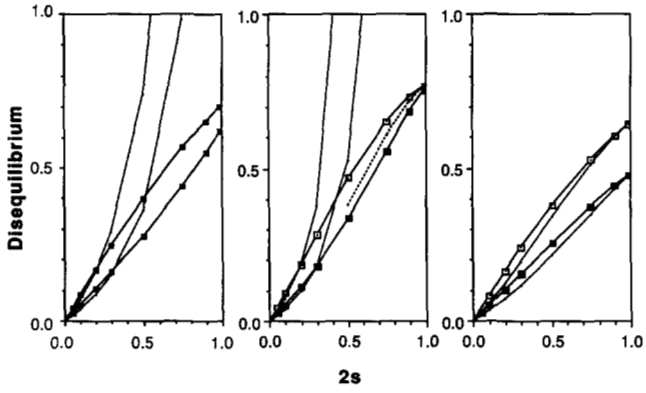

FIGURE 5.—Disequilibria. The maximum gametic correlation R is plotted against the per locus selection pressure $2s$. Lines represent values from simulations; all loci dominant (open squares), all loci codominant (closed squares), mean disequilibria when 1 locus is codominant, and 2 others are completely dominant (dashed lines in center graph). Dotted lines represent predictions based on Equation 9 and the actual dp/dx produced in the simulations (Figure 2). Left, 2 loci; center, 3 loci, using the standard order (reproduction–migration–selection); right, 3 loci using the order reproduction–selection–migration.

Thus the strength of natural selection in a hybrid zone can be measured simply by comparing the disequilibrium in the center of the zone with that produced in models of a particular selective process.

With a number of loci under weak selection and with no epistasis, the disequilibrium produced will be:

$$D_{AB} \approx \frac{\sigma^2}{r_{AB}} \frac{dp_A}{dx} \frac{dp_B}{dx} \qquad (9a)$$

(BARTON 1982, 1983) where r_{AB} is the recombination rate between loci A and B. With low s, the maximum gradient is $\sim k\sqrt{s/\sigma}$, and so disequilibrium should peak in the center of two identical clines, where:

$$R_{AB} = D_{AB}/0.5^2 \approx \frac{4sk^2}{r_{AB}}. \qquad (9b)$$

Simulations show that disequilibria do peak in the centers of clines (Figure 3), and depend almost linearly on selection pressure (Figure 5). When some loci are codominant and others are completely dominant, the average disequilibria are intermediate between the case where all loci are dominant and the case where all are codominant (Figure 5), as might be expected. The frequency-dependent selection used here causes a higher constant of proportionality than expected from Equation 9b (*i.e.*, $R_{AB}/s > 4k^2/r_{AB}$). This is for two reasons. Multiplicative fitnesses generate some nonadditive epistasis, which causes extra disequilibria. Adding more loci then increases the disequilibrium further, because the additional selection imposed by more disequilibria reduces cline widths.

Theory predicts that higher order disequilibria will also be produced in multilocus clines (BARTON 1983). For instance, third order disequilibria change from zero to positive values, switch to negative values, and then back to zero as we move across the clines (Figure 3), as expected (BARTON 1982).

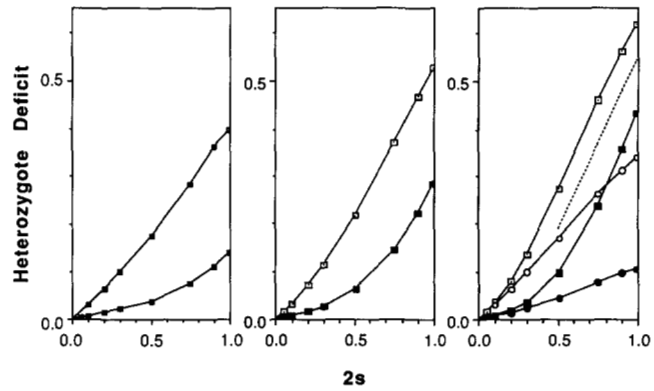

FIGURE 6.—Heterozygote deficit. The maximum heterozygote deficit F is plotted against $2s$. Dominant loci, open symbols; codominant loci, closed symbols. Left, 1 locus; center, 2 loci; right, 3 loci; dashed line indicates 1 codominant and 2 dominant loci. The standard order (reproduction–migration–selection) is plotted in square symbols; the order reproduction–selection–migration is plotted in circles.

Positive correlations between alleles within a locus are also expected because of gene flow. However, since Hardy-Weinberg equilibrium ($F_{IS} = 0$) results after one round of random mating, within-locus correlations, F_{IS}, are not expected to equilibrate at as high a level in the center of a cline as between-locus correlations, R; simulations confirm this (Figure 6).

Selection itself may cause deviations from Hardy-Weinberg. This is especially true for dominant loci. As we have already seen, a long tail of recessive alleles is produced in the edge of the cline where the recessives are rare. Here, recessive homozygotes will be produced from a large reservoir of heterozygotes, but they will be strongly selected against, producing a negative F_{IS}. By a similar argument, we expect selection to produce a positive F_{IS} in the other tail, where dominant phenotypes are rare, adding to that produced by gene flow. The effect of selection on F_{IS} is thus expected to be much lower on codominant loci than on dominant loci. In the simulations, F_{IS} is about three times as large in dominant clines than it is in codominant clines (Figure 3 and 6). In comparison, R (Figure 5) differs less than twofold between simulations with codominant and dominant loci, about as expected on the basis of differences in slope alone (Figure 2).

Epistasis: The assumption of multiplicative fitnesses implies that selection acts independently on each locus. In nature, color patterns under frequency-dependent selection may often be produced by epistatically interacting loci: for example, in Heliconius warning patterns (SHEPPARD *et al.* 1985; MALLET 1989), pure patterns are likely to be more memorable than hybrid patterns. A tension zone could potentially

TABLE 2

Effect of varying proportions of epistasis (e) and frequency-dependence ($2s$)

$2s$	e	$\sigma \times$ max. dp/dx	Speed/σ	Max. R	Max. F	Min. F
0.5	0	*A* 0.250	0.051	*AB/AC* 0.356	*A* 0.116	*A* 0.000
		B,C 0.314		*BC* 0.421	*B,C* 0.226	*B,C* −0.054
		Mean 0.292		Mean 0.378	Mean 0.190	
0.5	0.1	*A* 0.356	0.056	*AB/AC* 0.493	*A* 0.223	*A* 0.000
		B,C 0.368		*BC* 0.514	*B,C* 0.278	*B,C* −0.058
		Mean 0.364		Mean 0.500	Mean 0.256	
0.4	0.1	*A* 0.278	0.053	*AB/AC* 0.417	*A* 0.169	*A* 0.000
		B,C 0.292		*BC* 0.443	*B,C* 0.221	*B,C* −0.049
		Mean 0.287		Mean 0.425	Mean 0.203	
0.25	0.25	*A* 0.274	0.057	*AB/AC* 0.475	*A* 0.226	*A* 0.000
		B,C 0.261		*BC* 0.491	*B,C* 0.219	*B,C* −0.046
		Mean 0.266		Mean 0.481	Mean 0.221	
0	0.5	*A* 0.253	0.062	*AB/AC* 0.588	*A* 0.322	*A* 0.000
		B,C 0.226		*BC* 0.613	*B,C* 0.241	*B,C* −0.049
		Mean 0.235		Mean 0.597	Mean 0.190	

Results of simulations of three-locus tension zones with $\sigma^2 = 30$, using the standard order of processes: reproduction–migration–selection. Locus *A* is codominant, loci *B* and *C* are completely dominant, as in Figure 3.

be stabilized by such epistasis alone (BAZYKIN 1973), and we here model one such case.

Table 2 shows the effect of varying combinations of epistasic and frequency-dependent selection on a three-locus tension zone. If the clines are maintained purely by epistasis ($2s = 0$, $e = 0.5$), the shape and speed are little different from the pure frequency-dependent model ($2s = 0.5$, $e = 0$). The heterozygous deficit F is increased for codominant loci because the simple phenotype model of epistasis used here implies heterozygous disadvantage at codominant loci; dominant loci have similar F across models. Because epistasis favors particular gene combinations, disequilibria are strongly affected; purely epistatic selection causes correlations, R, in the center of the zone to increase by about 60% over those produced in a purely phenotypic model. However, if only a small proportion of the selection is due to epistasis (for example $2s = 0.5$, $e = 0.1$, compared with $2s = 0.5$, $e = 0$) the disequilibria are less strongly affected.

This analysis is somewhat crude because $2s$ and e are not really equivalent. However, Table 2 shows in a general way that large amounts of epistasis will lead to an overestimate of selection if disequilibria in a natural tension zone were assumed to result solely from frequency-dependent selection. With these reservations, it seems that the differences in the estimates will usually be less than twofold [see also BARTON (1983), p. 464].

Order of processes: The order of processes makes no difference at low s. The order assumed thus far has been reproduction-migration-selection. If s is high and the order is instead reproduction-selection-migration, the maximum gradients are reduced and become closer to the low s approximations (Figure 2); the speed is similar, but does not tail off at high s (Figure 4); disequilibria, R, are reduced and become closer to the small s approximation of equation 9 (Figure 5); and within-locus correlations, F, are reduced also (Figure 6). Many of these effects result because selection acts on genotypes in Hardy-Weinberg equilibrium, rather than on genotypes that have just been disturbed by migration. Additional effects are due to the partial reduction of disequilibrium by reproduction before selection. Because populations are nearer to linkage equilibrium, we expect that clines will not be attracted together as strongly. This prediction is borne out: for the model of Figure 3 and Table 2 under the order reproduction-selection-migration, codominant and dominant clines always separate for $2s \leq 1$, compared with clines under the standard order, in which the critical $2s \approx 0.5$.

DISCUSSION

Linear frequency-dependent selection on codominant genes produces, in the low selection limit, recursions that have the same form as those for heterozygote and disadvantage. Dominant clines are similar, but differ in that asymmetries of selection caused by the dominance produce an assymmetric cline. Dominant clines will also move, even in the absence of phenotypic selective differences: codominant clines do not move unless there are such selective differences.

Clines can provide estimates of gene flow, σ^2, and selection pressure, s, which do not require direct field measures (BARTON 1982; SZYMURA and BARTON 1986). For biologically possible values ($2s \leq 1$), these estimates will be correct to within a factor of ~2, whatever the exact form of selection. However, clines are often found in which only a few genes or chro-

mosomes strongly affect phenotypes, and presumably fitness [*e.g.*, Heliconius (MALLET 1986a, 1989; MALLET and BARTON 1989), *Bombus* (OWEN and PLOWRIGHT 1980), and birds (GILL 1973, 1980)]. To analyze tension zones, we will therefore often need to extend the weak selection theory, as is attempted here. These extensions are problematical because, whereas a number of models collapse into a single model at lows s, as we increase s the order of processes as well as other assumptions begin to matter. (Whether fitness differences at such strongly selected loci were built up gradually is not at issue here: provided $r \approx 0.01$ or less, intragenic recombinants are selected out more rapidly than they are produced by recombination in tension zones, so that they can be ignored for the purposes of inference).

Analytical theory is most useful if the order is reproduction-selection-migration. This might be partially true if the organism dispersed throughout its life, and selection occurred relatively early on. The abolition of F and reduction of R by mating and recombination just before selection make loci virtually independent for low numbers of unlinked loci. If there are many loci scattered around the genome the average r would be greatly reduced and disequilibria become important, requiring a different analysis (BARTON 1983).

On the other hand, in many species a burst of migration is followed by a relatively sedentary period during which selection is likely to occur. Then migration mainly precedes selection and the increasing deviations from Hardy-Weinberg under stronger selection cause narrower clines which in turn create stronger deviation from Hardy-Weinberg: this causes a strongly curvilinear relation between $\sqrt{s}$ and maximum gradient. This effect is magnified still further as the number of loci is increased because of a similar feedback effect on disequilibria (Figure 2). Under this order of processes simulations will be required to estimate σ^2 and s accurately from the cline data.

To make inferences about selection and gene flow we need measurements of at least two parameters from tension zones. The simplest feature to measure is width, defined $1/\text{max.}\ (dp/dx)$. This gives the ratio $\sigma/\sqrt{s}$ (Figure 2). One further measurement will fix the absolute values of each parameter. One possibility is to use heterozygote disadvantage F. However, F depends strongly on details of the model. F is more responsive to selection on dominant genes; unfortunately, heterozygote frequencies required for estimating F will usually be difficult to measure unless a locus is codominant in an alternative biochemical test, for example. The F at codominantly selected loci are much lower, and will be hard to detect, even for relatively high levels of selection (see Figure 6 and LEWONTIN and COCKERHAM 1959). However, measurements of F will be the only way to make inferences from single locus clines (*e.g.* MALLET 1986a).

Correlations between loci, R, equilibrate at a higher level than F because they are only partly broken down by recombination in each generation, will thus be easier to measure, and can give an efficient means of estimating the absolute value of s (Figure 5). Higher order disequilibria (Figure 3) could also be used, but require greater sample sizes for accurate estimation.

Finally, it should be borne in mind that the selection inferred by this means depends on a one parameter model of frequency-dependent selection. Consider warning color: among the actual parameters of importance are likely to be the epistasis between loci affecting the phenotype, the unpalatability of the aposematic species, relative numbers of predators and aposematic prey, one or more learning parameters of the predators, and the availability of alternative foods. The value of s that is inferred will give a good idea of the strength of selection, but not its exact nature. For reasonably large data sets, the accuracy of these estimates will be limited by uncertainties as to the model rather than by sampling error. Nonetheless, these measures are likely to be very useful given the inaccuracy of most other information on evolution: at the very least they can confirm the results of field experiments.

We thank the Natural Environmental Research Council for financial support, and N. SANDERSON, M. SLATKIN and an anonymous reviewer for comments on the manuscript.

LITERATURE CITED

BARTON, N. H., 1979 The dynamics of hybrid zones. Heredity **43:** 341–359.

BARTON, N. H., 1982 The structure of the hybrid zone in *Uroderma bilobatum*. Evolution **36:** 863–866.

BARTON, N. H., 1983 Multilocus clines. Evolution **37:** 454–471.

BARTON, N. H., and B. CHARLESWORTH, 1984 Genetic revolutions, founder effects, and speciation. Annu. Rev. Ecol. Syst. **15:** 133–164.

BARTON, N. H., and G. M. HEWITT. 1985 Analysis of hybrid zones. Annu. Rev. Ecol. Syst. **16:** 113–148.

BAZYKIN, A. D., 1969 Hypothetical mechanism of speciation. Evolution **23:** 685–687.

BAZYKIN, A. D., 1973 Population genetic analysis of disruptive and stabilizing selection. Part II. Systems of adjacent populations within a continuous area. Genetika **9:** 156–166.

BROWN, K. S., 1979 *Ecologia Geográfica e Evolucão nas Florestas Neotropicais.* Universidade Estadual de Campinas, Campinas, Brazil.

BROWN, K. S., P. M. SHEPPARD and J. R. G. TURNER, 1974 Quaternary refugia in tropical America: evidence from race formation in *Heliconius* butterflies. Proc. Roy. Soc. Lond. B **187:** 369–378.

ENDLER, J. A., 1977 *Geographic Variation, Speciation and Clines.* Princeton University Press, Princeton, N.J.

ENDLER, J. A., 1988 Frequency-dependent selection, crypsis, and aposematism. Phil. Trans. R. Soc. Lond. B **319:** 459–472.

FISHER, R. A., 1950 Gene frequencies in a cline determined by selection and diffusion. Biometrics **6:** 353–361.

GILBERT, L. E., and J. T. SMILEY, 1978 Determinants of local

diversity in phytophagous insects: host specialists in tropical environments, pp. 89–104 in *Diversity of Insect Faunas*, edited by L. A. MOUND and N. WALOFF. Blackwell, Oxford.

GILL, F. B., 1973 Intra-island variation in the mascarene white-eye *Zosterops borbonica*. Ornithol. Monogr. **12:** 1–66.

GILL, F. B., 1980 Historical aspects of hybridization between blue-winged and golden-winged warblers. Auk **97:** 1–18.

HALDANE, J. B. S., 1948 The theory of a cline. J. Genet. **48:** 277–284.

HEWITT, G. M., 1988 Hybrid zones—Natural laboratories for evolutionary studies. Trends Ecol. Evol. **3:** 158–167.

JOHNSON, M. S., J. MURRAY and B. CLARKE. 1987 Independence of genetic subdivision and variation for coil in *Partula sutaralis*. Heredity **58:** 307–313.

KIMURA, M., AND G. H. WEISS, 1964 The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics **49:** 561–576.

LANDE, R., 1982 Rapid origin of sexual isolation and character divergence in a cline. Evolution **36:** 213–223.

LANDE, R., 1985 The fixation of chromosomal rearrangements in a subdivided population with local extinction and recolonization. Heredity **54:** 323–332.

LEVIN, B. R., 1988 Frequency-dependent selection in bacterial populations. Phil. Trans. R. Soc. Lond. B. **319:** 459–472.

LEWONTIN, R. C., and C. C. COCKERHAM, 1959 The goodness-of-fit test for detecting natural selection in random-mating populations. Evolution **13:** 561–564.

MALLET, J., 1984 *Population Structure and Evolution in* Heliconius *Butterflies*. Ph.D. Dissertation, The University of Texas at Austin.

MALLET, J., 1986a Hybrid zones of *Heliconius* butterflies in Panama and the stability and movement of warning colour clines. Heredity **56:** 191–202.

MALLET, J., 1986b Gregarious roosting and home range in *Heliconius* butterflies. Natl. Geogr. Res. **2:** 198–215.

MALLET, J., 1986c Dispersal and gene flow in a butterfly with home range behavior *Heliconius erato* (Lepidoptera: Nymphalidae). Oecologia **68:** 210–217.

MALLET, J., 1989 The genetics of warning colour in hybrid zones of *Heliconius erato* and *H. melpomene*. Proc. R. Soc. Lond. B. **236:** 163–185.

MALLET, J., and N. H. BARTON, 1989 Strong natural selection in a warning color hybrid zone. Evolution **43:** 421–431.

MALLET, T. J., and M. C. SINGER, 1987 Individual selection, kin selection, and the shifting balance in the evolution of warning colours: the evidence from butterflies. Biol. J. Linn. Soc. **32:** 337–350.

MANI, G. S., 1985 Evolution of resistance in the presence of two insecticides. Genetics **109:** 761–783.

MAY, R. M., J. A. ENDLER, and R. E. MCMURTRIE, 1975 Gene frequency clines in the presence of selection opposed by gene flow. Am. Nat. **109:** 659–676.

MCLAIN, D. K., 1988 Male mating preferences and assortative mating in the soldier beetle. Evolution **42:** 729–735.

MOORE, W. S., 1979 A single locus mass-action model of assortative mating, with comments on the process of speciation. Heredity **42:** 173–186.

MOORE, W. S., 1987 Random mating in the Northern Flicker hybrid zone: implications for the evolution of bright and contrasting plumage patterns in birds. Evolution **41:** 539–546.

NAGYLAKI, T., 1975 Conditions for the existence of clines. Genetics **80:** 595–615.

OWEN, R. E., AND R. C. PLOWRIGHT, 1980 Abdominal pile color dimorphism in the bumble bee, *Bombus melanopygus*. J. Hered. **71:** 241–247.

ROUHANI, S., and N. H. BARTON, 1987 Speciation and the "shifting balance" in a continuous population. Theor. Pop. Biol. **31:** 465–492.

SHEPPARD, P. M., J. R. G. TURNER, K. S. BROWN, W. W. BENSON, and M. C. SINGER, 1985 Genetics and the evolution of muellerian mimicry in *Heliconius* butterflies. Phil. Trans. R. Soc. Lond. B. **308:** 433–613.

SLATKIN, M., 1972 On treating the chromosome as the unit of selection. Genetics **72:** 157–168.

SLATKIN, M., 1973 Gene flow and selection in a cline. Genetics **75:** 733–756.

SLATKIN, M., 1975 Gene flow and selection in a two-locus system. Genetics **81:** 209–222.

SLATKIN, M., 1981 Fixation probabilities and fixation times in a subdivided population. Evolution **35:** 477–488.

SZYMURA, J. M., AND N. H. BARTON, 1986 Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. variegata*, near Cracow in Southern Poland. Evolution **40:** 1141–1159.

WALSH, J. B., 1982 Rate of accumulation of reproductive isolation by chromosome rearrangements. Am. Nat. **120:** 510–532.

WRIGHT, S., 1932 The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proc. VIth Int. Congr. Genet. **1:** 356–366.

WRIGHT, S., 1969 *Evolution and the Genetics of Populations*. Vol. 2. *The Theory of Gene Frequencies*. University of Chicago Press, Chicago.

WRIGHT, S., 1980 Genic and organismic selection. Evolution **34:** 825–843.

WRIGHT, S., 1982 Character change, speciation, and the higher taxa. Evolution **36:** 427–443.

Communicating editor: M. TURELLI