Copyright © 1990 by the Genetics Society of America

Estimates of Selection and Gene Flow From Measures of Cline Width and Linkage Disequilibrium in Heliconius Hybrid Zones

James Mallet,* Nick Barton,† Gerardo Lamas M.,‡ José Santisteban C.,§ Manuel Muedas M.‖ and Harriet Eeley#

**Department of Entomology, Mississippi State University, Mississippi State, Mississippi 39762, †Galton Laboratory, Department of Genetics and Biometry, University College, London NW1 2HE, England, ‡Museo de Historia Natural, Universidad Nacional, Lima 14, Perú, §Academy of Natural Sciences, Philadelphia, Pennsylvania 19103, ‖Departamento de Biologia, Universidad Nacional de Trujillo, Trujillo, Perú, and #Department of Zoology, Cambridge University, Cambridge CB2 3EJ, England*

Manuscript received June 12, 1989
Accepted for publication December 18, 1989

ABSTRACT

Hybrid zones can yield estimates of natural selection and gene flow. The width of a cline in gene frequency is approximately proportional to gene flow (σ) divided by the square root of per-locus selection ($\sqrt{s}$). Gene flow also causes gametic correlations (linkage disequilibria) between genes that differ across hybrid zones. Correlations are stronger when the hybrid zone is narrow, and rise to a maximum roughly equal to s. Thus cline width and gametic correlations combine to give estimates of gene flow and selection. These indirect measures of σ and s are especially useful because they can be made from collections, and require no field experiments. The method was applied to hybrid zones between color pattern races in a pair of Peruvian *Heliconius* butterfly species. The species are Müllerian mimics of one another, and both show the same changes in warning color pattern across their respective hybrid zones. The expectations of cline width and gametic correlation were generated using simulations of clines stabilized by strong frequency-dependent selection. In the hybrid zone in *Heliconius erato*, clines at three major color pattern loci were between 8.5 and 10.2 km wide, and the pairwise gametic correlations peaked at $R \approx 0.35$. These measures suggest that $s \approx 0.23$ per locus, and that $\sigma \approx 2.6$ km. In *erato*, the shapes of the clines agreed with that expected on the basis of dominance. *Heliconius melpomene* has a nearly coincident hybrid zone. In this species, cline widths at four major color pattern loci varied between 11.7 and 13.4 km. Pairwise gametic correlations peaked near $R \approx 1.00$ for tightly linked genes, and at $R \approx 0.40$ for unlinked genes, giving $s \approx 0.25$ per locus and $\sigma \approx 3.7$ km. In *melpomene*, cline shapes did not perfectly fit theoretical shapes based on dominance; this deviation might be explained by long-distance migration and/or strong epistasis. Compared with *erato*, sample sizes in *melpomene* are lower and the genetics of its color patterns are less well understood. In spite of these problems, selection and gene flow are clearly of the same order of magnitude in the two species. The relatively high per locus selection coefficients agree with "major gene" theories for the evolution of Müllerian mimicry, but the genetic architecture of the color patterns does not. These results show that the genetics and evolution of mimicry are still only sketchily understood.

HYBRID zones and clines can provide estimates of gene flow and selection. The width of a cline, w, defined as the inverse of the maximum gradient of the cline, $1/(dp/dx)$, is approximately proportional to $\sigma/\sqrt{s}$ when selection is low (p denotes gene frequency, x denotes distance, σ is the gene flow, measured as the standard deviation of parent-offspring distances along one dimension, and s is the selection coefficient; weak selection, $s < 0.1$, is assumed). The constant of proportionality varies with the type of selection, but is usually between $\sqrt{3}$ and $\sqrt{12}$ for clines maintained by an ecological boundary, or stabilized by heterozygous disadvantage or frequency-dependent selection (e.g., ENDLER 1977; BARTON 1979a; MALLET 1986a; MALLET and BARTON 1989a).

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

This relationship enables one to estimate the intensity of selection acting in clines and hybrid zones from knowledge of cline width and gene flow. Gene flow has usually been estimated by means of mark-recapture experiments. Such studies have suggested that clines are often maintained by very low selection pressures, of the order of 10^{-5}–10^{-3} (ENDLER 1977; BARTON and HEWITT 1985). These selection pressures are so low as to be almost incredible: genetic drift and occasional population restructuring would often outweigh selection if such estimates were correct. In some cases, the low estimated selection has been used as evidence that clines are slowly decaying after secondary contact between races differing at effectively neutral loci (*e.g.*, ENDLER 1977; BARROWCLOUGH 1980). However, these ideas about selection depend strongly on the accuracy of the gene flow estimates. It has

recently been realized that measures of gene flow based on typical mark-recapture studies may often be underestimates (JONES *et al.* 1981; MOORE and DOLBEER 1989). Indirect estimates of gene flow based on geographic patterns of gene frequencies also tend to show that gene flow is more extensive than previously realized (SLATKIN 1987).

If genotypes at more than one locus can be detected in a set of clines, it becomes possible to estimate the selection and gene flow producing the clines without recourse to field experiments or other measures of dispersal. When individuals move between areas that differ in gene frequency at a number of loci, they carry with them gametes from their source population. Thus, gene flow (strictly, genotype flow or migration) causes correlations between loci within gametes. Gametic correlations are usually called linkage disequilibria, although correlated genes need not be linked, and an equilibrium of "disequilibrium" can be reached in a set of clines between the production of correlations by migration and their destruction by recombination. This steady-state disequilibrium is approximately proportional to dispersal divided by the widths of the clines (BARTON 1982):

$$D \approx \sigma^2/rw_1w_2$$

(where r is the rate of recombination between the two loci and w_1, w_2 are the widths of clines at these loci). As cline widths themselves depend on a ratio of migration to selection, the maximum D is directly proportional to s (MALLET and BARTON 1989a); if r is known, selection can be estimated by measuring disequilibria in a cline. So far, there are only two examples where linkage disequilibria have been used in the measurement of selection and migration (BARTON 1982; SZYMURA and BARTON 1986). This paper applies similar techniques to the analysis of hybrid zones in butterflies.

Heliconius butterflies have strongly differentiated color pattern races which are separated by hybrid zones of between 10 and 100 km wide, depending on the zone. The genetics of the color pattern differences are relatively easy to investigate in the laboratory (SHEPPARD *et al.* 1985), so that genotypes can be identified from field collections and gametic correlations can be estimated. Heliconius color patterns are often mimetic: some mimic other heliconiines (TURNER 1984; BROWN, SHEPPARD and TURNER 1974), others mimic ithomiine butterflies (BROWN and BENSON 1974). Heliconius are unpalatable to birds (BROWER, BROWER and COLLINS 1963); their bright color patterns almost certainly act as warning signals, and the mimicry is Müllerian. This imposes frequency-dependent selection against rare forms within a population (MALLET and SINGER 1987; ENDLER 1988), which can maintain clines (MALLET 1986a) similar to those caused by heterozygous disadvantage (BAZYKIN 1969; BARTON 1979a). Heliconius thus provide systems of multiple hybrid zones with good information on the genetics as well as on the nature of selection.

We have investigated hybrid zones between races of *Heliconius erato* and between those of its Müllerian comimic *Heliconius melpomene* in the Río Huallaga and Río Mayo drainages of Peru. The inheritance of the color pattern races of both species has been investigated (MALLET 1989). In this paper we present measurements of cline widths and disequilibria in the hybrid zones. We use the theory of clines maintained by frequency-dependent selection (MALLET and BARTON 1989a) to estimate levels of selection and gene flow from these data. We compare these indirect estimates with direct estimates of the same two parameters based on mark-recapture studies.

MATERIALS AND METHODS

Species studied and geographical location: *H. erato* was selected for study because there is good information on its genetics, ecology and behavior, and because this species is one of the commonest Heliconius in the field. Early mark-recapture work and observations of gregarious roosting sites seemed to indicate largely closed populations of Heliconius, with closely related individuals patrolling communal home ranges, and kin selection maintaining altruistic behavior. However, detailed studies revealed much less genetic structure than had been thought. Home ranges overlap, newly emerged individuals frequently disperse out of their parents' home ranges, and roosts seem to be little more than selfish groups of individuals attempting to avoid nocturnal dangers (MALLET 1986b, c). These observations of apparently continuous population structure tally with continuous clines of color pattern loci within hybrid zones (MALLET 1986a; this study).

We chose a hybrid zone between color pattern races of *erato* in the drainages of the Ríos Mayo and Huallaga near Tarapoto, Peru. Data were also gathered on its Müllerian comimic *H. melpomene*, which has a similar hybrid zone in the same place. In each species a "postman" race from the upper Huallaga and Mayo valleys hybridizes with a lowland "rayed" race from the Amazonian drainage of the lower Huallaga (Figure 1). The upland races have no comimics, but the lowland races belong to a mimicry ring consisting of up to about six other rayed heliconiine species in any one locality. However, other rayed comimics are usually rarer than *erato* and *melpomene*, whose color patterns may therefore control mimetic selection. This pair of hybrid zones was convenient for the following reasons: (1) The hybrid zones were between strongly divergent color patterns, so that one expects strong and detectable selection on warning color. However, hybrid zones between forms as divergent as those in Figure 1 are not uncommon in Heliconius (TURNER 1972; BROWN, SHEPARD and TURNER 1974). (2) The zone satisfied the requirement that two or more color pattern genes segregate, which is necessary for the measurement of gametic correlations between such genes. Based on the genetic survey of SHEPPARD *et al.* (1985), many hybrid zones between Amazonian and extra-Amazonian races of *erato* or *melpomene* satisfy these criteria. In contrast, zones within each subarea often separate races with only slight pattern differences due to fewer major loci (*e.g.*, MALLET 1986a). (3) Bureaucratic restrictions on biological research are minimal in Peru. (4) The sites had to be easily accessible.

FIGURE 1.—Heliconius from the hybrid zones. Top row: *H. erato favorinus* and *H. melpomene amaryllis*, from the Rio Mayo and upper Rio Huallaga. Bottom row: *H. erato emma* and *H. melpomene aglaope*, from the lower Rio Huallaga and Rio Ucayali drainages. Black is shown in black, yellow as clear, and red as stippled. Hybrids are abundant within a species, but interspecific hybrids are unknown; in fact the two species are in separate groups within the genus *Heliconius*.

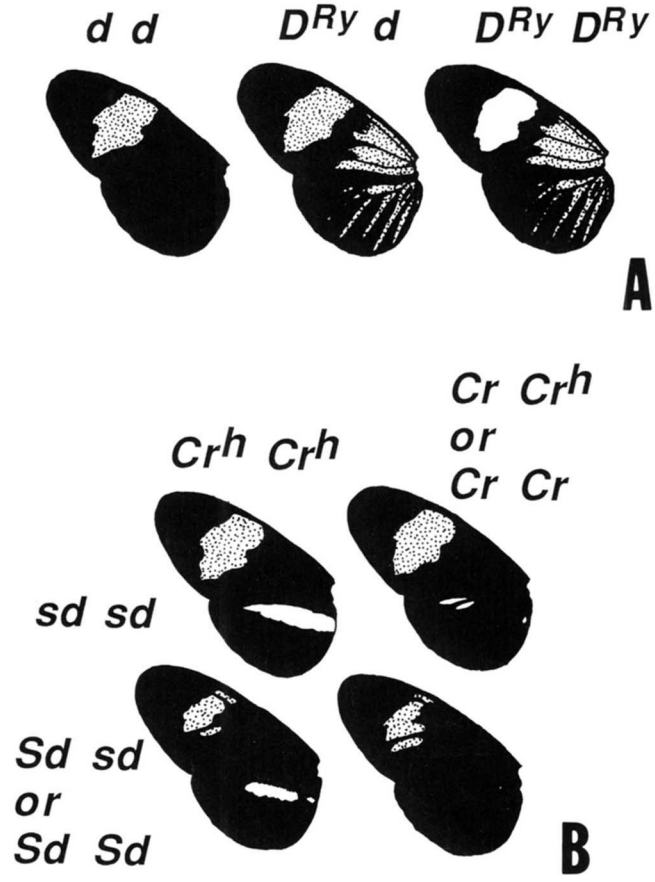

FIGURE 2.—The genetics of color pattern in *H. erato*. A, The effect of the D^{Ry} locus on the "dennis" (proximal red patch), ray, and forewing band color. B, The effect of the interaction between *Cr* and *Sd* on forewing band shape and hindwing yellow bar. Color scheme as in Figure 1.

In the Tarapoto zones, access was provided by the Tarapoto-Yurimaguas road, and by the navigability of the lower Río Huallaga. (5) In the region of these hybrid zones, vegetation is still largely natural, though there is small-scale agriculture in narrow strips along road and river margins. (6) The Tarapoto region of Peru was one of the few suitable sites in the eastern Andes of South America which is safe for foreigners.

Genetics of color patterns: In Heliconius, color pattern is often determined by relatively few loci of major effect (SHEPPARD *et al.* 1985), though there is evidence that some of the loci consist of tightly linked blocks, or "supergenes" (MALLET 1989). In Peruvian *erato*, a single locus D^{Ry} controls major elements of the color pattern: the D^{Ry} allele adds an orange-red patch on the forewing ("dennis") and orange-red hindwing "rays," and determines yellow color in the forewing band; the d^{ry} allele removes the dennis and ray, and determines a crimson-red forewing band. Heterozygotes are easily recognizable since the red components of the pattern are dominant: $D^{Ry}d^{ry}$ heterozygotes have ray and dennis, and a red forewing band in which some yellow scales may be present (Figure 2A; see also SHEPPARD *et al.* 1985). The *Sd* locus controls the shape of the forewing band: *sdsd* individuals have a broad forewing band, like that of the postman race; *Sdsd* and *SdSd* individuals are not clearly distinguishable and have a strongly narrowed forewing band. *Cr* interacts with *Sd* to remove the yellow hindwing bar, which is only present in *sdsd* Cr^hCr^h individuals. *Sd*− Cr^hCr^h individuals have a weak basal bar; *sdsd Cr*− individuals, recognizable by the broadness of their forewing bands, have hindwing yellow expressed only at the tips of their hindwing bars (Figure 2B). These loci are all unlinked.

In *melpomene* five genes controlling color pattern are known. A gene *N* interacts with another gene *B* to determine the color and shape of the forewing band (TURNER 1972; SHEPPARD *et al.* 1985), as shown in Figure 3A. There is now some evidence that a third locus, *M*, which is known to be unlinked to *N*, also affects the forewing band (MALLET 1989; J. MALLET and L. GILBERT unpublished results); however, further work is needed to distinguish *mm* N^BN^B from *M*-N^N- genotypes. The recessive allele *m* is Amazonian in origin, as is the similar though dominant N^N, and these two alleles appear to have similar frequencies in the hybrid zone. We expect some errors to be made by assuming, as is done here, that the yellow forewing bands are entirely due to N^N alleles, but we can do little better at present. In cases where this assumption leads to peculiar results, we show below that this is the effect expected when *mm* genotypes are ignored. A locus *Yb* controls the yellow hindwing bar, which is only produced in *ybyb* genotypes (Figure 3B). Although there is some penetrance of the yellow bar in *Ybyb* heterozygotes (SHEPPARD *et al.* 1985; MALLET 1989), for the purposes of estimating gene frequencies and disequilibria it is safest to interpret phenotypes at this and other *melpomene* loci as though they were completely dominant. Finally, a locus D^R controls only the presence or absence of dennis and ray (Figure 3C): the gene in *erato* is possibly homologous, see TURNER (1984), though the forewing band color is unaffected by D^R in *melpomene*, and there are major differences in the details of the hindwing ray pattern. These loci belong

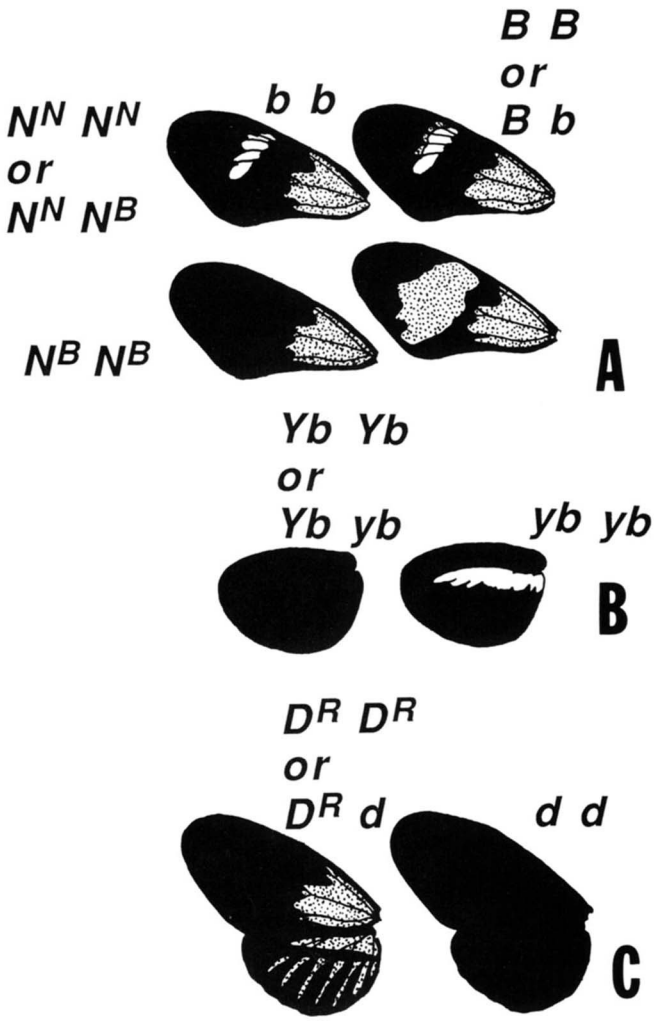

FIGURE 3.—The genetics of color pattern in *H. melpomene*. A, The effect of the interaction between *N* and *B* on forewing band shape and color. The *B* allele has a much smaller effect on an N^N-background than on an N^BN^B background. The locus *M* (not shown) can affect this interaction. Homozygous *mm* produce a phenotype similar to N^N- *B*- in an insect that is genotypically N^BN^B *B*-. B, The effect of *Yb* on the yellow hindwing bar. C, The effect of D^R on the dennis and ray patterns. Color scheme as in Figure 1.

to two linkage groups: *N* and *Yb* are linked with a recombination frequency in males of about 1%; *B* and D^R are linked with a recombination frequency in males of 27% (SHEPPARD et al. 1985; as modified by MALLET 1989, and J. MALLET and L. GILBERT, in preparation). *M* is not linked to *N* (J. MALLET and L. GILBERT, in preparation), and there is weak evidence from the field (see RESULTS) that *M* is linked to *B* and D^R. Recombination between linked genes is unknown in Heliconius females (TURNER and SHEPPARD 1975), which, like other female Lepidoptera have achiasmatic oogenesis (SUOMALAINEN, COOK and TURNER 1973), so the average recombination rates are about half the values given above for males.

A complete account of the genetics, interactions, linkage, and variability of penetrance for these hybridizing taxa is given in MALLET (1989).

Sampling: Butterflies were sampled in the adult stage from a number of sites near Tarapoto. Each site was located on a USAF high flight aerial photo composite, and the grid position was determined to within the nearest 0.5 km (Table 5 in the APPENDIX). Most of the sites (see map in MALLET and BARTON 1989b) were located near a line stretching NNE; this transect was used to estimate cline width since the hybrid zones crossed this transect approximately at right angles (Table 5 in the APPENDIX). The butterflies were caught at flowers, or flying along streams or roads. High flying individuals of both sexes were often brought to within net range by rapidly waved crimson rags (BROWN and BENSON 1974). A total of 1572 *erato* and 903 *melpomene* from 53 and 51 sites respectively were scored for phenotype in the analysis (Tables 6 and 7 in the APPENDIX).

Approach taken in the analysis: Clines may vary along a transect in two basic parameters: width and position. In addition, clines may be steep in the center but have longer "tails" at the edges than an idealized single locus cline, which is likely if a cline at one locus is trapped in a set of clines at other loci (BARTON 1983). Dominance or asymmetric selection will cause a cline to be asymmetric, with recessive (or less selected) alleles forming a long tail in the end of the cline where dominant (or heavily selected) alleles are common (HALDANE 1948; MALLET and BARTON 1989a).

In principle all these parameters–position, width, number of loci, and dominance–might be estimated from the data. However, because the number of loci and dominance are already known under the assumption of selection for warning color, we take these as part of the model used to estimate cline shape and position. There is good evidence that the color patterns of Heliconius warn predators, and that this should lead to frequency-dependent selection against the rare form. We have found no allele frequency differences at 38 soluble enzyme loci between the races of either species (J. MALLET and P. KING, in preparation; see also TURNER, JOHNSON and EANES 1979), and no evidence for hybrid inviability (MALLET 1989; see also SHEPPARD *et al.* 1985). Therefore, it seems likely that the loci determining color pattern are heavily selected, and that there is not a strong polygenic barrier to gene flow across the zone. Even if other kinds of selection on other loci are important, the conclusions reached will still hold approximately, providing that the loss in mean fitness due to the background loci is not too great (BARTON 1986). There is direct field evidence that selection on these loci is strong ($s \geq 0.1$, MALLET and BARTON 1989b, which agrees with our indirect estimates here), so analytical theory would be inaccurate if used to predict cline shape. We therefore use simulations of clines at three loci under frequency dependent selection to give the theoretical expectations (MALLET and BARTON 1989a).

Ideally, clines of genotype frequencies could be fitted to the data, giving information on dominance, gene frequencies, heterozygote deficits, and disequilibria in a single step. In practice this would be difficult, since the theoretical expectations for each set of parameters must be generated by rather lengthy simulations. In any case an appropriate algorithm was not available: the construction of such an algorithm would have been problematic since dominant clines tend to move, and clines differing in dominance tend to move apart (MALLET and BARTON 1989a). Geographic discontinuities and epistatic selection may prevent cline movement in nature (see the DISCUSSION). Instead, we fitted cline shape for each locus separately to gain estimates of $\sigma/\sqrt{s}$. Disequilibria were then estimated in polymorphic populations for which there were sufficient data, which enabled estimates to be made of both *s* and σ. It should be noted that, because there are strong disequilibria between the loci, estimates of the shapes of clines are not independent. For this reason, we use only the average widths of the clines within a species to estimate $\sigma/\sqrt{s}$.

Estimation of heterozygote deficit: Heterozygotes can

be reliably detected at only one locus, D^{Ry} in *erato*. For this locus, the heterozygote deficit, F, was estimated for each polymorphic sample using likelihood, and the results were combined across sites by summing the log likelihoods for different values of F in each site.

Estimation of cline shape: To fit the clines in *erato* we used three-locus simulations in which one locus was codominant (to fit D^{Ry}) and two other loci were dominant (to fit *Cr* and *Sd*). The per locus fitnesses (see Equation 1 below) were assumed equal and to combine multiplicatively across loci. Multiplicative fitnesses seem more realistic, but in any case produce similar results to additive models, providing that selection is not too high. Migration was modeled as a binomial diffusion process (approximately Gaussian) between discrete, linearly arranged demes. Binomial migration can be "tuned" so that the demes approximate a continuous population, as appears to occur in Heliconius. A selection pressure of $s = 0.25$ per locus was used for the simulated clines because this was the lowest selection pressure that prevented the dominant clines moving away from the codominant cline (for details, see MALLET and BARTON 1989a). Strong selection is only one reason why clines within a species may be concordant in nature; other possibilities include partial dispersal barriers or epistasis. As we shall see, the observed disequilibria are also consistent with high selection pressures, for which we have direct evidence (MALLET and BARTON 1989b). The order of processes was assumed to be eclosion, migration, selection, followed by mating and egg-laying—this order mimics the order of processes in Heliconius where dispersal precedes the establishment of home range and reproduction (MALLET 1986b, c). These simulations were completely deterministic, and have been presented and fully discussed elsewhere (MALLET and BARTON 1989a).

In *melpomene* five color pattern loci require modeling. Two, *N* and *M*, are at present impossible to separate phenotypically; we are forced to ignore the locus *M*. Two loci, *N* and *Yb* are so tightly linked that the problem can be further simplified by treating them as the same locus. Thus effectively 3 loci need be modeled, considerably reducing the complexity of the simulations used to fit the clines. D^R (Amazonian allele dominant) and *B* (Amazonian allele recessive) are linked with an average recombination rate of 13.5%. The third locus, *N* + *Yb*, is not linked to the other two. These clines were fitted to a simulation model similar to that used for *erato* with three loci: the first (dominant) was linked to the second (recessive) with $r = 0.135$, and a third (dominant) locus simulated was unlinked. The loci had opposing dominance, which tends to cause opposing cline movement (MALLET and BARTON 1989a); in this case a per-locus selection of at least $s = 0.3$ was required to prevent the clines from moving apart.

The phenotypic data of Tables 6 and 7 in the APPENDIX were then fitted to the simulation results by means of likelihood, one locus at a time. For the dominant loci, this entailed assuming Hardy-Weinberg equilibrium ($F = 0$) within populations. This assumption conflicts with theoretical predictions for F: under the selection estimated below for *erato* ($s \approx 0.23$), the simulations predict that heterozygote deficits in the dominant clines will peak at $F \approx 0.25$ at dominant loci. The assumption that $F = 0$ would give a maximum error for gene frequency estimation of about 8% (e.g. if $p = 0.7$, then the estimated $p = \sqrt{(q^2 + Fpq)} \approx 0.62$. This would tend to give a position biased towards the dominant end of the cline, by a maximum of about $0.06w$ (position is measured at the point of $p = 0.5$: here, $0.06w \approx 0.6$ km). There would be rather less error in the width, w, itself. Estimates of cline width w and position x_0 were obtained for each locus. We also tested whether the actual shapes of clines at individual loci fitted simulated loci with the actual, visual dominance better than simulated loci with alternative levels of dominance. In performing such tests, field gene frequencies were estimated using the actual dominance of the loci as observed in the crosses. Only the dominance of the theoretical cline was varied. Genotypic frequencies from the simulations, which depend on cline position and width, were used as predictors of the field phenotypic frequencies. The fit was judged by differences in $\log_e$ likelihood, also called "support" (ΔL). Support limits, the limits where the $\log_e$ likelihood drops to two units below the maximum ($\Delta L \approx 2.0$), were estimated for each parameter assuming the other parameter to be fixed at its most likely value. These values mark points on the support surface, in this case a two-dimensional ellipsoid of support (see Figure 3 in MALLET and BARTON 1989b, for examples of such ellipsoids). Support limits along any one axis, as given here, correspond approximately to 95% confidence limits in large samples (EDWARDS 1972). Similarly, when two fits have $\Delta L > 2.00$, the most likely fit is said to be "significantly better" than the less likely fit.

In defense of the gametic correlation coefficient: Maximum and minimum values, both of D and of standardizations of D such as the correlation coefficient $R = D/\sqrt{(p_1q_1p_2q_2)}$ used here, are all dependent on the gene frequencies (HEDRICK 1987). This has caused a certain amount of despair that disequilibria can ever be compared between samples (LEWONTIN 1988). The correlation coefficient is perhaps the most useful standardized value since it is easily understood, being already used in other statistical fields. With fixed gene frequencies, R cannot usually vary throughout the range $(-1, +1)$; for example if $p_1 = p_2 = p$, the range is $-\min(q/p, p/q) < R < 1$. This might seem a drawback. However, note that WRIGHT's inbreeding coefficient F_{IS}, another correlation coefficient whose use has not been contested, also "suffers" from an inability to vary between -1 and $+1$. A standardized coefficient of disequilibrium like R, which only reaches 100% of maximum when $p_1 = p_2$, may be rather useful: some "recombinants" must be present in the sample if $R < 1$. Suppose tight linkage, together with strong selection or drift, tended to produce strong positive or negative disequilibria. These factors would also alter the gene frequencies–disequilibria cannot be considered in isolation, with the gene frequencies fixed. Finally, an alternative standardization recommended by HEDRICK, LEWONTIN's $D' = D/D_{max}$ is not suitable here because D_{max} changes at $D = 0$.

Estimation of gametic correlations: Gametic correlations (linkage disequilibria) were estimated for each polymorphic sample of more than 15 individuals for each species, using likelihood. Because other workers may disagree about methods of estimation, we include the numbers of individual phenotypes in the APPENDIX Table 6 (*erato*) and APPENDIX Table 7 (*melpomene*). Essentially, the method of HILL (1974) was used, though the maxima and inflections of likelihood were found by trial and error aided by computer calculations of the likelihoods (i.e. probabilities of observed phenotypic distributions given particular gene frequencies, deviations from Hardy-Weinberg, and two-way and three-way disequilibria). As presented in the tables, the disequilibria represent most likely values, with support limits as described above shown in parentheses. The support limits were calculated contingent on Hardy-Weinberg equilibrium ($F = 0$: estimates of disequilibria are unaffected by this procedure), maximum-likelihood gene frequencies, and higher order disequilibria at their joint most likely values.

In *erato*, all loci were unlinked, so there is no *a priori*

expectation that either gene frequencies or pairwise disequilibria should be heterogeneous within samples. Therefore the single most likely D was estimated over all three pairs of loci, assuming equal gene frequencies and equal disequilibria across loci at each site. In 5 samples on the Northern (Amazonian) end of the cline, however, gene frequencies were significantly heterogeneous (see Table 3). For these samples gene frequencies were estimated separately, and the support curve for each pairwise D was produced, transformed to a support curve of the correlation coefficient, R, and then added to the two other support surfaces to give the support curve for the single best R value over the three loci. There were no significant heterogeneities in the value of R within any site. This method therefore allows a single best pairwise correlation to be estimated using information from all three loci in each sample.

In *melpomene*, it is harder to combine information across loci. Six pairwise correlations can be estimated. Two of these are between linked genes (D^R-*B* and *N*-*Yb*), and so are expected to be greater than disequilibria between unlinked genes. Correlations between unlinked genes and N are complicated by the effect of the *mm* genotype, whose effects cannot be distinguished from that of N^N-. We therefore estimate these separately from correlations between unlinked genes not involving N. Finally, the correlations between unlinked genes D^R and *Yb* may be similar to those between *B* and *Yb*, but cannot be combined to give the overall correlation between unlinked genes since *B* and D^R are linked, giving non-independent estimates for the two disequilibria. For *melpomene*, then, it becomes extremely cumbersome to estimate all of the gene frequencies, the pairwise, the three-way, and the four-way disequilibria, and in any case the results are liable to error because of the difficulty of interpreting the phenotypes. Instead, each pairwise disequilibrium and its support limits have been estimated separately as a rough guide, contingent on $F \approx 0$ and maximum likelihood gene frequencies.

Estimation of selection and gene flow: The estimates of cline width and disequilibria were then used to estimate the strengths of selection and gene flow. The estimated widths of gene frequency clines give an estimate of the migration:selection ratio ($\sigma/\sqrt{s}$). Selection is then estimated independently from gene flow using the gametic correlations, and gene flow is estimated by solving for the width. Since frequency-dependent selection is thought to maintain these clines, we use results developed especially for this case (MALLET and BARTON 1989a). We here define s in the simplest model of weak frequency-dependent selection as equivalent to that in weak heterozygote disadvantage. The model used was as follows (MALLET and BARTON 1989a, Equation 4):

$$W_{AA} = 1 - 2s[f_{aa} + (1 - h)f_{Aa}]$$

$$W_{aa} = 1 - 2s(f_{AA} + h f_{Aa})$$

$$W_{Aa} = 1 - 2s[h f_{aa} + (1 - h)f_{AA}]$$

(h measures the dominance, and f_{AA}, f_{Aa} and f_{aa} are the frequencies of the genotypes). This definition of s aids comparisons with clines maintained by heterozygous disadvantage, but leads to the peculiarity that s in frequency-dependent selection can never be greater than 0.5 to avoid negative fitnesses of homozygotes. The "s" used in a previous study (MALLET 1986a), and measured in a mark-recapture experiment (MALLET and BARTON 1989b) was thus in fact $2s$, as defined here.

The uncertainties in the analysis depend as much on confidence in the model assumptions (see DISCUSSION) as on sample size, so the estimates of selection and migration are given without support limits. Calculation of support limits would anyway be difficult without an extensive restructuring of the methods to estimate all parameters simultaneously, requiring the development of an appropriate algorithm (see above under *Approach taken in the analysis*). However, we feel the present analysis is robust. Estimates of selection and gene flow are unlikely to be more than a factor of about two from the real value (BARTON 1983; MALLET and BARTON 1989a).

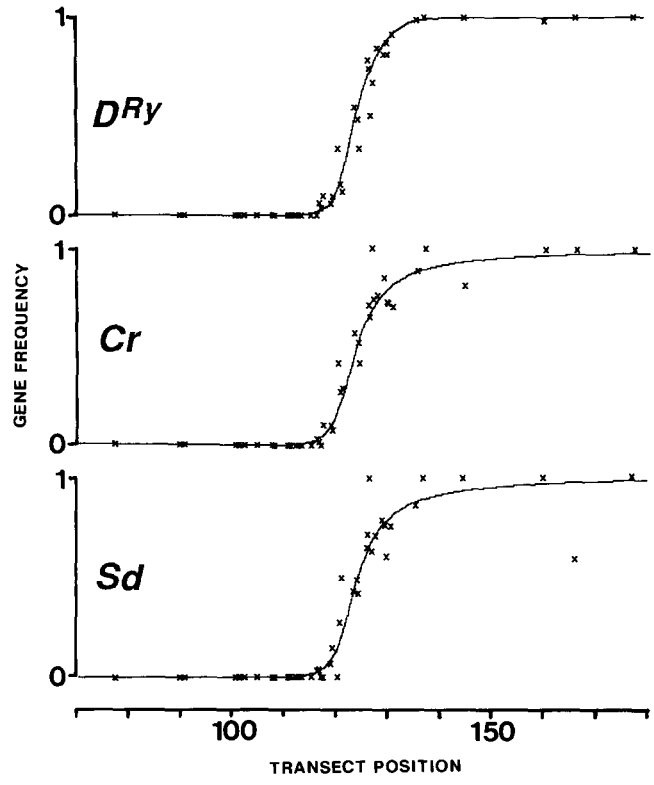

FIGURE 4.—Color pattern clines in *H. erato*. Estimated gene frequencies at each site for all three color pattern loci are shown as points, together with the best fitting theoretical clines for each locus as given in Table 1.

RESULTS

Heliconius erato: The primary data are given in Table 6 of the APPENDIX. Heterozygotes are distinguishable at only one of the loci, D^{Ry}. At no site was F significantly different from zero at this locus. Maximizing the likelihood over all sites, $F = 0.01$ (support limits −0.016, 0.095), again not significantly different from zero. The strong asymmetry of the support interval is due to the impossibility of a strongly negative F when gene frequencies are extreme, which occurs in some samples near the edges of the hybrid zone.

Genotypic frequencies in Table 6 of the APPENDIX were fitted to the simulated clines; the best fits are shown in Figure 4. Estimated cline widths and positions are given in Table 1, and their support limits are given in Table 2. The actual shape of each cline was most similar to the simulated cline with the same

TABLE 1

Positions and Widths (in km) of Heliconius clines

Locus	Parameter	Recessive	Codominant	Dominant
H. erato				
D^{Ry}	Position	125.11	*123.57*[a]	123.25
	Width	7.71	*8.46*[a]	6.67
	Log likelihood	−681.89	−583.50	−597.57
Cr	Position	125.91	124.20	*123.74*[a]
	Width	9.10	11.78	*10.16*[a]
	Log likelihood	−402.92	−289.78	−280.93
Sd	Position	126.18	124.48	*124.17*[a]
	Width	7.98	11.53	*10.15*[a]
	Log likelihood	−411.13	−306.16	−289.43
H. melpomene				
D^{R}	Position	122.53	122.18[a]	*122.18*
	Width	10.18	11.74[a]	*11.81*
	Log likelihood	−204.50	−190.67	−192.16
N	Position	123.08	122.50[a]	*122.45*
	Width	11.50	11.67[a]	*11.69*
	Log likelihood	−216.17	−187.53	−190.19
Yb	Position	123.54	123.02[a]	*123.02*
	Width	11.52	11.95[a]	*12.00*
	Log likelihood	−220.95	−194.03	−196.26
b	Position	*121.65*	121.33	121.30[a]
	Width	*19.39*	16.02	13.34[a]
	Log likelihood	−254.73	−249.16	−248.37

The actual data were fitted to stabilized theoretical clines generated in three-locus simulations. For *H. erato*, the simulations had $s = 0.25$, $\sigma = 30$, and consisted of one codominant locus and two completely dominant loci, with all three loci unlinked, exactly as in Figure 3 of MALLET and BARTON (1989). To obtain an approximate theoretical expectation for a recessive cline, a dominant cline from this same simulation was simply inverted. For *H. melpomene* a similar three-locus model was used, in which one gene was recessive and linked with $r = 0.135$ to a second, dominant gene (to simulate the *b*- D^{R} gene pair); the third gene was unlinked to the other two and dominant (to simulate the very tightly linked *N-Yb* gene pair). The simulations were run at $s = 0.3$ and $\sigma = 30$. Theoretical clines for codominant loci were the same as those used to fit *erato* data.

Italics: expected on the basis of simple visual dominance to give the best fit.

[a] These estimates are used because they give the actual best fit.

TABLE 2

Support limits for cline parameters

Species	Locus	Width/km (limits)	Position/km (limits)
H. erato	D^{Ry}	8.46 (7.77 9.20)	123.57 (123.22 123.92)
	Cr	10.16 (8.88 11.51)	123.74 (123.21 124.30)
	Sd	10.15 (8.82 11.52)	124.17 (123.62 124.69)
H. melpomene	D^{R}	11.74 (10.07 13.60)	122.18 (121.57 122.77)
	N	11.67 (10.11 13.38)	122.50 (121.91 123.11)
	Yb	12.00 (10.38 13.71)	123.02 (122.41 123.61)
	B	13.34 (10.76 16.72)	121.30 (120.50 122.11)

dominance (Table 1). Thus D^{Ry} fitted significantly better to a nearly symmetrical codominant cline than to asymmetric dominant ($\Delta L = 14.06$) or recessive clines ($\Delta L = 98.39$), as expected from its actual visual codominance ("dominance" of a cline is used here to mean that the allele found in the Amazon basin is dominant). The two actually dominant loci, *Cr* and *Sd*, fit significantly better to simulated dominant loci than to codominant loci ($\Delta L = 8.84$ and 16.73, respectively) or recessive loci ($\Delta L = 121.31$ and 121.69). The heterogeneity of gene frequencies at different loci within sites where Amazonian alleles are common (Table 3) may be explained similarly: the allele d^{ry} is significantly rarer than the other Mayo alleles Cr^{h} and *sd* in the Amazonian end of the hybrid zone, presumably because its codominance allows it to be detected by predators even when heterozygous. As well as being more symmetrical, the D^{Ry} cline is also significantly narrower than those at the other two loci, though not by much (Table 2). This narrowness fits with the expected stronger selection: D^{Ry} changes the color patterns of both forewing and hindwing more extensively than do *Cr* or *Sd* (Figure 2).

Estimated disequilibria are given in Table 3. In *erato* all loci are unlinked. Thus there is no expectation for strong differences between pairs of loci, and in-

TABLE 3

Gametic associations in *H. erato*

Site	Site no.	*N*	Best *q*	G_2 (heterog.*q*)	Best *D*	G_2 (heterog.*R*)	Best *R* (limits)
km 40–41	28 + 29	42	0.03	0.304	−0.001	0.00	−0.03 (−0.03[a], 0.19)
km 46	30	31	0.08	0.692	0.020	5.68	0.27 (0.04, 0.50)
km 48	31	26	0.11	1.192	0.044	3.68	0.45 (0.22, 0.64)
km 54	33	19	0.23	1.942	0.058	0.86	0.33 (0.06, 0.53)
km 58	34	84	0.50	0.886	0.050	0.12	0.20 (0.05, 0.34)
bel Pongo	36	15	0.68	0.294	−0.040	3.21	−0.18 (−0.37, 0.27)
Sh'jilla	40	22	0.52	1.298	0.086	2.70	0.19 (−0.13, 0.48)
Yumbatos	42	74	0.73	3.980	0.037	0.34	0.19 (−0.01, 0.40)
km 62, S	43	25	0.71	0.598	0.066	0.60	0.32 (−0.05, 0.65)
km 62, N	44	144	0.79	9.296**		5.19	0.24 (0.04, 0.46)
km 63	45	45	0.81	0.372		3.68	0.36 (0.01, 0.67)
km 68A	48	127	0.80	15.600***		1.13	0.18 (−0.01, 0.42)
km 68B	49	69	0.84	14.978***		1.91	0.48 (0.37, 0.48[a])
km 72	50	160	0.96	20.476***		1.54	0.18 (−0.03[a], 0.31[a])

* G significant at $P < 0.05$.
** G significant at $P < 0.01$.
*** G significant at $P < 0.001$.
[a] Support limit is at the boundary of possible gamete space Best *q*, best *D*, and best *R* indicate the most likely gene frequencies, disequilibria, and gametic correlations estimated jointly over all three genes. G_2 is the G value for heterogeneity, with two degrees of freedom, of the indicated estimates.

deed, no heterogeneity of gametic correlation was found within any site. We therefore assume that the three pairwise correlations are equal in each site, which also helps to make efficient use of the samples. Of the 14 sites for which *R* was estimated, seven are significantly different from 0, and in all seven the correlation was positive. Only two samples gave a negative estimate of *R*.

In theory, the gametic correlations are expected to peak in the center of the hybrid zone (BARTON 1982; MALLET and BARTON 1989a), but the pattern is not noticeable from the data, presumably because of the small sample sizes, and possibly also because disequilibria might vary between sites with local differences in migration or selection. Thus we can get only a crude idea of the peak correlation. If we exclude the two sites with *p* outside the range 0.1–0.9 (km 40–41 Tarapoto-Yurimaguas and Davidcillo) the average *R* is about 0.27; however, the peak may be much higher, as much as 0.45–0.48; we will take 0.35 as a sensible value. Comparing this value with the appropriate set of results in Figure 5B of MALLET and BARTON (1989a), we find that this peak disequilibrium corresponds to a selection pressure of $s \approx 0.23$ per locus. We use Figure 2C of MALLET and BARTON (1989a) to infer that this level of selection implies a standardized maximum slope, σ/w, of 0.27. The observed average width is about 9.59 km. Thus $\sigma \approx 0.27 \times 9.59 \approx 2.6$ km.

As we have seen, migration produces interlocus correlations, but migration should also produce deviations from Hardy-Weinberg. This expectation has been quantified for clines under frequency-dependent selection (MALLET and BARTON 1989a). Much higher levels of deviation (F) should be produced at dominant loci than at codominant loci, because selection itself accentuates the deviation in the former case. As we have seen, the estimate of F has wide support limits in the single locus, D^{Ry} in *erato*, available for testing; for $s \approx 0.23$ in the simulated codominant locus we expect $F \approx 0.08$ (MALLET and BARTON 1989a, Figure 6), which is within the support limits of the observed F. It is unfortunate that we cannot estimate F at the dominant loci: based on the models we would expect F to peak at ~0.25. Such high levels of F, if they exist, will mean that the gene frequency estimates, which assume $F = 0$, are subject to a maximum error of about 6% (see under *Estimation of cline shape* above), giving a maximum error of about 0.6 km in cline position and width. It would be an interesting test of the model to obtain estimates of F at dominant genes by breeding wild-caught individuals.

***Heliconius melpomene*:** The primary data are given in Table 7 of the APPENDIX. The best fit to each of the clines is shown in Table 1 and Figure 5. The clines coincide approximately, with centers between km 121 and 123 of the transect, and have widths as shown in Table 1. Support limits for position and width are given in Table 2. The D^R-cline fits better to the codominant cline in the simulations used to fit *erato* than it does to the dominant linked gene in the *melpomene* simulations, but not significantly ($\Delta L = 1.49$). Each of the loci *N* and *Yb* fit significantly better to codominant clines than to dominant unlinked clines of the appropriate simulations, though only marginally ($\Delta L = 2.65$ and 2.22, respectively). These three

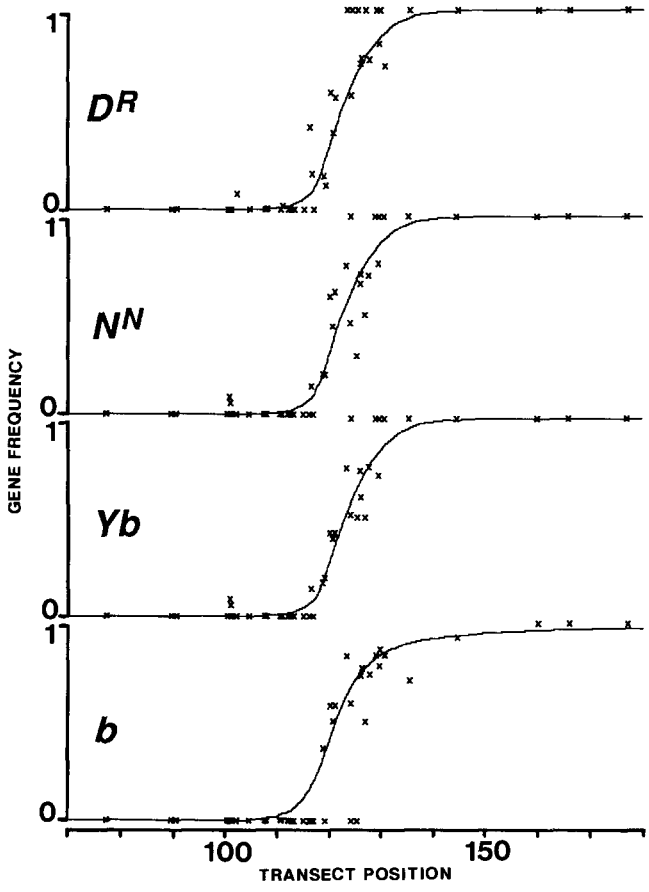

FIGURE 5.—Color pattern clines in *H. melpomene*. Estimated gene frequencies at each site for four color pattern loci are shown as points, together with the best fitting color pattern clines for each locus as given in Table 2.

loci fit far worse to recessive clines than to dominant clines (ΔL = 12.35, 25.98 and 24.69, respectively). Locus *B* fits better to a dominant cline than to either codominant or recessive clines (ΔL = 0.79 and 6.36, respectively). Since *B* is "recessive" in the sense that the recessive *b* allele is characteristic of the Amazonian race, it might be expected to fit better to the recessive locus in the simulations. In fact *B* fits much better to a dominant cline than to a recessive cline, the reverse of what we expect.

It is clear from Figure 5 that there is far more scatter around the expectations in *melpomene* than in *erato*, so part of the lack of fit could be due to greater genetic drift and/or smaller sample sizes in *melpomene*. The better fit of *D*, *N*, and *Yb* to codominant than to recessive clines is due to a higher than expected frequency of D^R, N^N, and *Yb* alleles in a very few Huallaga and Mayo populations (Figure 5). This pattern could be caused by occasional long-distance migrants, whereas Gaussian migration was assumed in the simulations. This explanation is supported by the high correlation between *N* and *Yb* (R = 1) between Shapaja and Chazuta (Table 4) where both are rare.

Another reason for the poor fit to simulated loci

TABLE 4

Gametic associations in *H. melpomene*

Site no.	*N*	Best *p*				Best *R* (support limits) for each gene pair					
		D^R	N^N	*Yb*	*b*	$D^R - b$	$N^N - Yb$	$Yb - b$	$N^N - b$	$D^R - Yb$	$D^R - N^N$
7–10	63	0.02	0.02	0.02	0		1.00 (0.56, 1.00)			−0.02 (−0.02, 0.45)	−0.02 (−0.02, 0.45)
28–29	18	0.22	0.12	0.12	0		1.00 (0.72, 1.00)			0.13 (−0.19, 0.55)	0.13 (−0.19, 0.55)
30	22	0.20	0.20	0.17	0.37	0.65 (0.11, 0.65)	0.72 (0.34, 0.89)	0.02 (−0.32, 0.60)	0.65 (0.11, 0.65)	0.32 (−0.14, 0.68)	0.44 (−0.01, 0.76)
31	17	0.13	0.20	0.20	0		1.00 (0.78, 1.00)			0.18 (−0.18, 0.62)	0.18 (−0.18, 0.62)
33	48	0.39	0.44	0.39	0.50	0.79 (0.39, 0.79)	0.89 (0.77, 0.89)	0.42 (−0.04, 0.79)	0.88 (0.34, 0.88)	0.34 (0.00, 0.61)	0.47 (0.16, 0.70)
34	17	0.58	0.46	0.51	0.59	0.97 (−0.11, 0.97)	0.89 (0.56, 0.89)	0.19 (−0.44, 0.85)	0.33 (−0.34, 0.75)	0.87 (0.49, 0.87)	0.78 (0.38, 0.78)
39	27	0.57	0.62	0.39	0.58	0.98 (0.06, 0.98)	0.63 (0.31, 0.63)	0.66 (0.03, 0.68)	0.92 (−0.03, 0.92)	0.30 (−0.24, 0.61)	0.70 (0.24, 0.89)
40	16	1.00	0.75	0.75	0.83		1.00 (0.38, 1.00)	0.78 (−0.15, 0.78)	0.78 (−0.15, 0.78)		
42	43	0.74	0.66	0.74	0.73	0.98 (0.07, 0.98)	0.83 (0.53, 0.83)	0.21 (−0.24, 0.78)	0.46 (−0.08, 0.84)	0.43 (−0.14, 0.82)	0.30 (−0.23, 0.71)
43	70	0.76	0.71	0.60	0.77	0.98 (0.09, 0.98)	0.79 (0.61, 0.79)	0.65 (0.19, 0.68)	0.86 (0.33, 0.86)	0.36 (−0.07, 0.62)	0.51 (−0.06, 0.79)
44	33	0.75	0.70	0.75	0.74	0.96 (−0.05, 0.96)	0.86 (0.45, 0.86)	0.04 (−0.32, 0.71)	0.24 (−0.26, 0.88)	0.61 (−0.04, 0.93)	0.51 (−0.12, 0.84)
48–49	49	0.80	0.80	0.75	0.80	0.99 (0.01, 0.99)	0.87 (0.48, 0.87)	−0.03 (−0.26, 0.55)	0.11 (−0.23, 0.75)	0.54 (−0.05, 0.85)	0.63 (0.02, 0.94)

Best *p* and best *R* indicate most likely estimates for each parameter.

with matching dominance could be the greater genetic interaction in *melpomene*: the forewing band in this species is affected by three genes: *N*, *B*, and *M* (Figure 3). *M* has been ignored for the moment, but it could be linked to *B* (see below). These interactions apparently lead to a greater interdependence between loci (*i.e.*, epistasis, disequilibria) than in *erato* or the simulated clines. The allele *B* causes a dominant addition of red color to the forewing; however, the area of red added is far greater in N^BN^B genotypes than in N^N- genotypes (Figure 3A). Thus the presence or absence of red is likely to be more highly selected in the Mayo end of the cline where N^B is common than in the Amazonian end where N^N is common. We would therefore expect the Mayo end of the cline in *B* to be steeper than the Amazonian end, the reverse of that expected from a consideration of dominance alone. We have not attempted simulations of such complexity. In addition, locus *B* is linked in repulsion to D^R, which causes further interactions–this association was included in the simulations. The difficulties caused by these interactions in *melpomene* make inferences for this species less reliable than for *erato*.

Disequilibria between the *melpomene* genes are given in Table 4. Correlations between linked genes (D^R-*B* and *N*-*Yb*) were high, always above 0.6, and all except two cases were at the theoretical maximum given the gene frequencies in the samples. Correlations between *N* and *Yb* have probably been reduced somewhat because the effect of locus *M* was not taken into consideration. Of 27 "recombinants" between *N* and *Yb* detected in these samples (see Table 7 in the APPENDIX) of the hybrid zone, 21, a fraction of 0.78 (0.59, 0.91), are of the apparent type N^N- *ybyb*, though 0.50 is the fraction expected: only 6 are N^BN^B *Yb*-. This is readily explained. Most of the N^N- *ybyb* phenotypes are probably N^BN^B *ybyb mm* genotypes, the *mm* homozygotes giving the appearance of N^N- genotypes. Thus we expect the correlations between *N* and *Yb* to have been underestimated; they must in reality be much closer to 1.00.

We cannot use disequilibria at linked genes to estimate selection since we have not allowed linkage to vary from 50% in models used to produce expectations (MALLET and BARTON 1989a), though we expect that disequilibria between linked genes to be stronger than those between unlinked genes, as observed (Table 4). Ignoring the first two sites (Shapaja-Chazuta, and km 40–41 Tarapoto-Yurimaguas) which are on the edges of the hybrid zone, the average correlations between unlinked genes are as follows: *Yb*-*B*, 0.29; *N*-*B*, 0.60; D^R-*Yb*, 0.42; and D^R-*N*, 0.49. Many of the sites have support limits that include zero, so the estimates are not reliable. However, nearly all of these correlations are positive (39/40). The correlations apparently involving *N* are highest, possibly because of linkage between *B*, D^R and *M*. This would mean *M B*, *m b*, *m* D^R and *M* d^r gametes would occur more frequently than if unlinked, and there would be an apparent association between the allele N^N and the *b* D^R chromosome because of the similarity of *mm* and N^N- genotypes. If *M* is assumed to be on the *B* D^R chromosome, it would seem to be closer to *B* than D^R, leading to the lower apparent association between D^R and *N* than between *B* and *N*. This may be reading too much into the data: the differences in association between unlinked genes could also have been produced by sampling variance, random drift, differences in selection, and/or epistasis.

Ignoring the correlations involving *N*, we are left with only two reliable pairwise correlations between unlinked genes, both involving *Yb*. The average correlation for these is 0.36 (again excluding the two sites at the edge of the hybrid zones), but the peak may be somewhat more; we shall take 0.40 as a sensible guess for the peak *R*.

This disequilibrium between unlinked genes in *melpomene* suggests selection similar to that found in *erato*, perhaps a little more (about 0.25, interpolating from Figure 5B of MALLET and BARTON 1989a). The average width of clines is slightly greater than in *erato*, 12.30 km, leading to a greater estimated migration rate, $\sigma \approx 0.30 \times 12.30 \approx 3.7$ km (interpolating from Figure 2C of MALLET and BARTON 1989a). Although more dubious than that for *erato*, the data for *melpomene* do suggest that selection and migration are of the same order in the two species.

DISCUSSION

Validity of the assumptions: Confidence in these results, that $s \approx 0.2 - 0.3$, and migration, $\sigma \approx 2 - 4$ km generation$^{-1/2}$ in both species, will depend on one's belief in the assumptions of the model. The effects of a number of these assumptions are discussed below:

Selection is frequency-dependent: We have assumed that clines are stabilized solely by frequency-dependent selection on the color pattern genes. It has been shown above that the shapes of clines in *erato* match those expected on the basis of the visual dominance of the pattern that each locus produces. The same is not true for *melpomene*, but this poorer fit could be explained by greater tendency to genetic drift, smaller sample sizes, or the greater gene interactions and uncertainty of the inheritance in *melpomene*. It should also be pointed out that the frequency-dependence used in the models is linear; real frequency-dependent selection due to predators is almost certainly nonlinear (MALLET 1986a; MALLET and BARTON 1989a). The present methods measure the "effective" selection that would be required to give similar clines if the frequency-dependence were linear.

These methods are relatively robust to the form of

selection. Using the result of BARTON (1982) cited in the Introduction to analyze the hybrid zone of *erato* as though it were due to weak heterozygote disadvantage, $\sigma^2 \approx Drw_1w_2 \approx 0.0875 \times 0.5 \times (9.59)^2$ km^2/generation; here, $\sigma \approx 2.01$ km. Since, under heterozygote disadvantage, $w \approx \sqrt{8}\,\sigma/\sqrt{s}$ (BAZYKIN 1969); $s \approx 0.77$. Selection is then greater than, though of the same order of magnitude as that estimated assuming frequency dependence. However, the greater estimate of selection is due more to the failure of analytical models when s is high, than to the type of selection. If selection were estimated based on models of strong heterozygote disadvantage, the results would be more similar to those for frequency dependence.

Selection is equal across loci: We have not investigated the effect of varying the selection on different loci in the model, although it would be surprising if selection were exactly equal across loci in nature. But the selection pressures are unlikely to be very different since all the loci analyzed had major phenotypic effects, and their cline widths are similar. Since the selection was estimated from an average of the pairwise disequilibria, it is an average measure of s.

The selection parameter can be used to compare across models: The selection measured here is defined by the frequency-dependent models of MALLET and BARTON (1989a). In these models, clines at loci with heterozygous disadvantage were shown to be similar to those at codominant loci under frequency-dependent selection, except that selection against heterozygotes was about 1.5 times as effective for the same value of s. We here define s by reference to heterozygote disadvantage (see Equation 1 in MATERIALS AND METHODS; MALLET and BARTON 1989a), which aids comparisons with clines maintained by underdominance. However, even changing the dominance of the locus under selection will tend to change the effect of the parameter s in frequency-dependent selection: judging by the reduction in width, a dominant locus in a cline experiences about 1.5 times as much selection as a codominant locus in a cline under the same s (MALLET and BARTON 1989a). This illustrates the fact that s can only be defined with respect to a particular model: comparisons between models will be approximate. Of course, general surveys of the strength of natural selection, such as that by ENDLER (1986), will be liable to the same problem.

Order of processes: The order of selection, migration, and reproduction makes little difference at low selection pressures where many models conform to the diffusion approximation, but can considerably affect the results at high selection pressures (MALLET and BARTON 1989a). The model used here assumed that selection occurred immediately after eclosion from the pupa, and was followed by dispersal and then by mating and reproduction. *H. erato* differs slightly from this ideal. Males eclose, and then disperse during a reproductive refractory period lasting a few days, set up home ranges, and mate in these new home ranges. Females are usually mated at or very soon after eclosion, before dispersal. Females then disperse to set up a home range before laying eggs, and rarely mate twice (MALLET 1986b). Thus the actual pattern of dispersal has the effect of creating more disequilibrium than under the simple model because the female carries not only her own gametes but also the gametes of her mate. This is equivalent to her offspring migrating *in utero*. The greater disequilibrium produced would cause the dispersal of adults, σ, to have been overestimated by a maximum of about $\sqrt{2}$. However, the σ measured here will be an approximately correct measure of actual gene flow.

Processes occur in sequence: Heliconius are continuous-brooded, but the models used were in discrete time. At low selection pressures, discrete generations conform to the diffusion approximation. At high selection, approximate spatial continuity can be ensured by binomial migration (MALLET and BARTON 1989a); but whether discrete time affects the results is unclear. Arguably there would be little effect because the groups of foreign migrants appearing in the model's discrete time are treated similarly in the models to the way predators are expected to accumulate experience with a number of single migrants in continuous time.

Migration is Gaussian: To mimic continuous populations in the discrete deme simulations, dispersal was modeled by a binomial (approximately Gaussian) distribution. However, many studies have shown that gene flow is highly leptokurtic (BATEMAN 1951; DOBZHANSKY and WRIGHT 1943; ENDLER 1977); this is also true for *erato* (reanalysis of data in MALLET 1986b). We could imagine that such leptokurtosis was produced by occasional long-distance migrants that distributed themselves uniformly over the whole area. This would lead to high correlations beyond the edges of the hybrid zone, even though the individuals that produce this disequilibrium might be rare. Considerations of this sort might explain why there are no apparent peaks of disequilibria in the centers of the hybrid zones (Tables 3 and 4), as expected under Gaussian migration (Figure 3 of MALLET and BARTON 1989a). Simulations show that leptokurtosis affects the peak correlations at the center of the hybrid zone very little: outside the zone there will be a moderate level of correlation caused by the long-distance migrants alone. However, there are other possible reasons for the lack of strong peaks in the disequilibria, not least of which are the small sample sizes.

Ignoring epistasis: In the models, the per-locus fitnesses were combined multiplicatively, giving some epistasis especially at high selection intensities. However, the genetics of the color patterns are also highly

epistatic in a different sense: pattern elements such as the yellow hindwing bar and the forewing band may depend on a number of genes for their complete expression (Figures 2 and 3). If this epistasis has a strong effect on fitness, selection might increase the observed disequilibria to make the selection and migration estimates too high. However, the effect is unlikely to be strong if, as we suspect, the selection is mainly produced by predators which memorize unpalatable prey phenotypes, making the relative abundance of different prey patterns paramount (MALLET and SINGER 1987). Epistasis will cause only a minor additional increase in mortality for multilocus hybrid patterns.

It should be noted that clines can be maintained by epistasis alone, and we have performed some simulations to investigate partial epistatic selection (MALLET and BARTON 1989a) in which "pure" AABB and aabb genotypes were assumed fitter than AaBB, AaBb, AAbb, and other "hybrid" genotypes. Very approximately, epistatic selection of this form will produce up to twice as much gametic correlation as the equivalent level of frequency-dependent selection (MALLET and BARTON 1989a). Thus the actual level of selection is unlikely to be less than half as much as that estimated here, even if the selection were largely epistatic.

Transect runs at right angles through the hybrid zones: The transect was chosen for easy access, and so may not cross the hybrid zone perpendicularly. This would cause the width of the hybrid zone to be overestimated (though not by much if the difference from right angles is less than about 20°), while leaving disequilibria unaffected. This would cause migration to be slightly overestimated.

Predators assumed stationary: In the model, predators are represented merely as selection pressures that depend on the current gene frequencies in the population. Any movement of real predators will cause them to select incorrectly, at least initially, to local butterfly gene frequencies; this "selection flow" will increase the width of the cline, but will hardly alter the disequilibria, which depend largely on butterfly dispersal. The most likely predators, (birds, especially jacamars) probably disperse chiefly as juveniles because most have home ranges and can live for 10 yr or more. Since birds learn prey patterns quickly (BROWER *et al.* 1963), it is probable that juvenile dispersal is of minor importance for a butterfly with a 3 month life cycle. Most of the local birds will be entrained on the local butterflies. The size of the home range is likely to be of more importance. Jacamars (Galbulidae), which are likely predators (BENSON 1972; CHAI 1986; MALLET and BARTON 1989b), have home ranges about 500–1000 m across (CHAI, personal communication), giving a radial σ (standard deviation of movement about the center of the home range) of up to about 300 m. This would not affect the width of a 10-km cline appreciably.

Cline position and potential movement: The hybrid zones within each species appear to be maintained by selection of about the minimum required to prevent the clines from moving apart because of differences in dominance. Epistasis might also cause these clines to stay tightly together: a hypothetical pure race consisting of homozygous recombinants should be less memorable than actual races because its pattern would be fuzzier and less striking (MALLET 1989). The zones in *erato* and *melpomene* are likely to stay together because of mimicry, a form of interspecific frequency-dependent selection. Coincidence of clines within and between species may also be aided by ecological gradients or density troughs which prevent movement, as discussed below for whole hybrid zones.

Hybrid zones with dominance or selective asymmetries can move. With $s \approx 0.23$ and $\sigma \approx 2.6$ km, we expect the *erato* zone to move at about $0.05 \times 2.6 = 0.13$ km/generation (see Figure 4 of MALLET and BARTON 1989a) towards the Mayo valley due to dominance alone. Assuming four generations a year, this could cause substantial hybrid zone movement in historical time, about 50 km per century. Instead, the Heliconius hybrid zones seem well established on the eastern slopes of the eastern Andes. (However, poor collecting data means there is no convincing evidence that they have not moved 50 km since they were first collected, mostly around the turn of this century). Moving clines are easily trapped by slight barriers to dispersal, which could exist because of random differences in density, or where ecological conditions are unsuitable (BARTON 1979a, b; HEWITT 1988). The area near Pongo de Cainarache where the hybrid zones occur is at the base of the easternmost Andes, and receives the highest rainfall in the region. Possibly this rainfall makes butterfly activity more difficult, and the area acts as a genetic sink which traps the hybrid zone. Alternatively, each race may be adapted to different ecological conditions (BENSON 1982; MALLET and BARTON 1989b). It will be interesting to resample the area in 100 years.

Comparison with field estimates: The ultimate test of these estimates of selection and migration based on theory is to compare them with direct experimental estimates. The problem of estimating gene flow from mark-recapture experiments has already been mentioned in the introduction. It is perhaps no surprise that a study of dispersal in *erato* gave a field-estimated σ only 1/9 as large (MALLET 1986b) as the present estimate of 2.6 km. The discrepancy could be explained in many ways; a combination of different dispersal in different areas (the dispersal experiment was performed in Costa Rica), the possibility of individuals moving outside the recapture area (which was

only about 2 km × 1.5 km), and uncertainty in the models used here seem among the most likely causes.

Selection in a field introduction of *erato* with foreign color patterns across this hybrid zone was about $s \approx 0.11$ per locus, assuming multiplicative fitnesses, but with a wide margin of error [MALLET and BARTON 1989b: in that paper, "s," here $2s$ (see MATERIALS AND METHODS), was given as 0.17 per locus, with the calculation being based on additive fitnesses; more realistic multiplicative fitnesses, as used in the simulations, give $2s \approx 0.22$, so the appropriate $s \approx 0.11$]. However, even this measure is difficult to interpret as a clear result of predation on foreign color patterns (MALLET and BARTON 1989b). In spite of these difficulties, the field estimate is similar, though somewhat lower, than the estimate in this paper of $s \approx 0.23$.

Cynics may find these discrepancies quite large. Nevertheless the level of agreement is heartening considering the many assumptions that must be made, both in the present analysis and especially in interpreting the field experiments. As a general rule, if we find narrow coincident clines with strong disequilibria, the total selection will be of the same order as the gametic correlations, and dispersal will be a substantial fraction of the cline width. If only a few selected loci are involved, each locus will suffer a substantial fraction of the total selection, as here. Heliconius are mimetic, giving visual evidence for selection, but it was unknown whether mimetic selection coefficients were 10^{-5} or 90%–see MALLET (1986a). The disequilibria measured in this study provide convincing evidence that mimetic selection in both species of Heliconius is more than 10% per locus with major phenotypic effects.

Implications for the understanding of mimicry: It is thought that the evolution of Müllerian mimicry may require a major initial mutation which brings a color pattern close to that of the model. Polygenic modification could then perfect the mimicry. Mutations with small effects alone would not work: they should be selected against because they would neither remind predators of the model nor remain protected by their own species' warning colors (NICHOLSON 1927; TURNER 1984; SHEPPARD *et al.* 1985). The relatively strong per-locus selection measured in this study is broadly consistent with this theory.

However, the genetics of the color patterns are not entirely consistent with the NICHOLSON-TURNER-SHEPPARD model. One problem is that there is very little evidence for polygenic modification of color pattern loci. Epistasis is common, but can mostly be attributed to the major loci themselves or elements closely linked to them (SHEPPARD *et al.* 1985; MALLET 1989). Some of these putative genes do appear to consist of several very tightly linked loci (MALLET 1989) which might suggest that modifiers have accumulated near the major genes. However, the evolution of these "supergenes" is unexpected in Heliconius. In polymorphic Batesian mimicry, selection against intermediates may lead to the evolution of closer linkage (CLARKE and SHEPPARD 1960). In Müllerian mimicry polymorphisms are selected against, and are transient except where mean fitness is lowest–in hybrid zones: without stable polymorphisms increased linkage is not expected to evolve. An alternative explanation (which we do not find altogether satisfactory) both for the existence of supergenes in Müllerian mimicry, and for the lack of unlinked modifiers, is that there are simply few chromosomal regions which can affect the color pattern in Heliconius (MALLET 1989). Whatever the true explanation, it is clear that there are still discoveries to be made about a textbook example of adaptation, the evolution of mimicry and warning color.

We thank the Natural Environmental Research Council, the Royal Society, the Nuffield Foundation, CONCYTEC, and Mrs. G. W. BORLASE for financial support, and the people of San Martín for their generous hospitality. We are very grateful to S. D. KNAPP, who helped by maintaining our sanity and rearing larvae. We are also grateful to an anonymous reviewer, A. W. PORTER, J. C. SCHNEIDER, M. TURELLI and C. E. WATSON for helpful comments on the manuscript. This paper was approved for publication as journal article no. J-7255 of the Mississippi Agricultural and Forestry Experiment Station, Mississippi State University, project no. MIS-2122.

LITERATURE CITED

BARROWCLOUGH, G. F., 1980 Genetic and phenotypic differentiation in a wood warbler (genus *Dendroica*) hybrid zone. Auk **97:** 655–668.

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

BARTON, N. H., 1979b Gene flow past a cline. Heredity **43:** 333–339.

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., 1986 The effects of linkage and density-dependent regulation on gene flow. Heredity **57:** 415–421.

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

BATEMAN, A. J., 1951 Is gene dispersion normal? Heredity **4:** 253–263.

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

BENSON, W. W., 1972 Natural selection for Müllerian mimicry in *Heliconius erato* in Costa Rica. Science **176:** 936–939.

BENSON, W. W., 1982 Alternative models for infrageneric diversification in the humid tropics: tests with passion vine butterflies, pp. 608–640 in *Biological Diversification in the Tropics,* edited by G. T. PRANCE. Columbia University Press, New York.

BROWER, L. P., J. VAN Z. BROWER and C. T. COLLINS, 1963 Experimental studies of mimicry. 7. Relative palatability and Müllerian mimicry among Neotropical butterflies of the subfamily Heliconiinae. Zoologica **48:** 65–84.

BROWN, K. S., and W. W. BENSON, 1974 Adaptive polymorphism associated with multiple Müllerian mimicry in *Heliconius numata* (Lepidoptera: Nymphalidase). Biotropica **6:** 205–228.

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. R. Soc. Lond. B **187:** 369–378.

CHAI, P., 1986 Responses of jacamars to butterflies. Biol. J. Linn. Soc. **29:** 161–189.

CLARKE, C. A., and P. M. SHEPPARD, 1960 Supergenes and mimicry. Heredity **14:** 175–185.

DOBZHANSKY, T., and S. WRIGHT, 1943 Genetics of natural populations. XV. Rate of diffusion of a mutant gene through a population of *Drosophila pseudoobscura*. Genetics **32:** 303–324.

EDWARDS, A. W. F., 1972 *Likelihood*. Cambridge University Press, Cambridge, U.K.

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

ENDLER, J. A., 1986 *Natural Selection in the Wild*. Princeton University Press, Princeton, N.J.

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

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

HEDRICK, P. W., 1987 Gametic disequilibrium measures: proceed with caution. Genetics **117:** 331–341.

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

HILL, W. G., 1974 Estimation of linkage disequilibrium in randomly mating populations. Heredity **33:** 229–239.

JONES, J. S., S. H. BRYANT, R. C. LEWONTIN, J. A. MOORE and T. PROUT, 1981 Gene flow and the geographical distribution of a molecular polymorphism in *Drosophila pseudoobscura*. Genetics **98:** 157–178.

LEWONTIN, R. C., 1988 On measures of gametic disequilibrium. Genetics **120:** 849–852.

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

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

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

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

MALLET, J., and N. H. BARTON, 1989a Inference from clines stabilized by frequency-dependent selection. Genetics **122:** 967–976.

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

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

MOORE, W. S., and R. A. DOLBEER, 1989 The use of banding recovery data to estimate dispersal rates and gene flow in avian species: case studies in the red-winged blackbird and common grackle. Condor **91:** 242–253.

NICHOLSON, A. J., 1927 A new theory of mimicry in insects. Aust. Zool. **5:** 10–104.

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. Philos. Trans. R. Soc. Lond. **308:** 433–613.

SLATKIN, M., 1987 Gene flow and the geographic structure of populations. Science **236:** 787–792.

SUOMALAINEN, E., L. M. COOK and J. R. G. TURNER, 1973 Achiasmatic oogenesis in the heliconiine butterflies. Hereditas **74:** 302–304.

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.

TURNER, J. R. G., 1972 The genetics of some polymorphic forms of the butterflies *Heliconius melpomene* (Linnaeus) and *H. erato* (Linnaeus). II. The hybridization of subspecies of *H. melpomene* from Surinam and Trinidad. Zoologica **56:** 125–157.

TURNER, J. R. G., 1984 Mimicry: the palatability spectrum and its consequences, pp. 141–161 in *The Biology of Butterflies* edited by R. I. VANE-WRIGHT and P. R. ACKERY. Academic Press, London.

TURNER, J. R. G., M. S. JOHNSON and W. F. EANES, 1979 Contrasted modes of evolution in the same genome: allozymes and adaptive change in *Heliconius*. Proc. Natl. Acad. Sci. USA **76:** 1924–1928.

TURNER, J. R. G., and P. M. SHEPPARD, 1975 Absence of crossing-over in female butterflies (*Heliconius*). Heredity **34:** 265–269.

Communicating editor: M. TURELLI

APPENDIX

Coordinates for hybrid zone data are given in Table 5, genotypic data for *H. erato* are given in Table 6, and *H. melpomene* genotypic frequency data are given in Table 7.

TABLE 5

Coordinates for hybrid zone data

Site No.	Site Name	km N	km E	Transect position
1	Above Pachiza	9189.0	303.0	0.51
2	Sacanche	9217.5	313.0	30.39
3	Near Buenos Aires	9250.0	352.0	77.31
4	Quebrada Mamonaquihua	9268.0	346.5	90.68
5	Achinamiza	9292.5	405.0	
6	Pongo de Aguirre	9290.5	402.5	
7	Hills above Chazuta	9273.5	372.5	107.66
8	km 15–17 Shapaja-Chazuta	9269.5	368.5	102.25
9	km 12–14 Shapaja-Chazuta	9269.0	367.0	101.11
9a	km 7–10 Shapaja-Chazuta	9270.5	364.0	101.04
10	km 1–5 Shapaja-Chazuta	9272.0	361.5	101.21
11	Cuñumbuque-Sisa	9276.5	328.5	89.82
12	Quebrada Vainilla	9291.0	324.0	100.55
13	San Antonio de Cumbaza	9292.0	344.0	110.75
14	km 9–13 Rodríguez de Mendoza-Omia	9297.0	238.0	
15	Río Jebil	9293.0	243.5	
16	Río Serranoyacu	9369.5	226.5	
17	54 km W. Rioja	9357.5	257.5	
18	Tarapoto (nr. airport)	9278.5	349.0	101.14
19	Tarapoto (Río Shilcayo)	9285.5	350.5	108.03
20	Tarapoto (Río Ahuashiyacu)	9280.5	353.0	104.77
21	km 7 T-Y	9283.5	354.0	107.89
22	km 15 T-Y, Cataratas de Ahuashiyacu	9286.5	355.0	111.01
23	km 17 T-Y, La Antena	9287.0	356.5	112.15
24	km 22 T-Y	9288.0	356.0	112.80
25	km 26 T-Y	9289.5	354.0	113.20
26	km 28 T-Y, Río Ponasillo	9292.0	353.5	115.18
27	km 30 T-Y, Río Tiraco	9293.0	355.5	117.00
28	km 40 T-Y, El Paredón	9289.5	360.5	116.23
29	km 41 T-Y, La Carachamera	9289.5	361.5	116.69
30	km 46–46.5 T-Y	9293.0	359.5	118.86
31	km 47.4–47.7 T-Y	9294.0	358.5	119.28
32	km 49–52 T-Y	9295.0	358.5	120.16
33	km 54 T-Y	9296.5	357.0	120.79
34	km 58 T-Y, above Pongo de Cainarache	9300.0	357.5	124.12
35	Hills S of Pongo de Cainarache	9300.5	357.0	124.33
36	Río Cainarache below Pongo	9301.5	360.5	126.85
37	Shucshuyacu, below Pongo de Cainarache	9301.0	360.5	
38	Cerro Isco	9296.5	349.5	
39	Shitariyacu, above Shapajilla	9299.0	353.0	121.14
40	Yuracyacu, above Shapajilla	9301.0	354.0	123.38
41	Near Yumbatos, S of Río Yuracyacu	9302.0	356.5	125.43
42	Near Yumbatos, N of Río Yuracyacu	9303.0	355.5	125.84
43	km 62 T-Y, S of Río Yuracyacu	9302.0	358.0	126.12
44	km 62 T-Y, N of Río Yuracyacu	9303.0	358.5	127.68
45	km 63 T-Y, New Road to Barranquitas	9304.0	360.0	128.83
46	km 66 T-Y, Tioyacu	9305.5	359.0	129.69
47	San Miguel de Shanusi	9318.5	346.0	
48	km 68A T-Y, Convento, stns. 1–11	9307.0	355.5	129.38
49	km 68B T-Y, Convento, stns. 12–13	9308.0	356.0	130.50
50	km 72 T-Y, Davidcillo	9310.5	361.5	135.28
51	km 75 T-Y, Bonilla	9313.0	360.0	136.79
52	km 82 T-Y, Naranjal	9322.0	359.0	144.29
53	km 25–26 Yurimaguas-Tarapoto	9338.0	362.0	159.84
54	km 17 Yurimaguas-Tarapoto	9340.5	370.0	165.78
55	Yurimaguas, right bank of Río Huallaga	9348.0	379.5	176.84

Coordinates for the sites are based on the Peruvian grid system, as used on U.S. Air Force high altitude aerial photographic maps. These coordinates can be converted to latitude and longitude by noting that, near the equator, 1′ of latitude or longitude ≈ 1.85 km. As a reference point, the airport at Tarapoto is at 76°22′W, 6°31′S, and (9279.6 km N, 348.3 km E) on the grid.

The transect was chosen to cut the hybrid zone approximately at right angles, and to be near the majority of the sites. The transect chosen stretched between (9190 km N, 300 km E) and (9380 km N, 400 km E) on the Peruvian grid system. Transect positions were calculated by dropping a perpendicular from each site onto this transect line, and calculating the distance along the transect segment from (9190 km N, 300 km E) to the site of intersection with the perpendicular.

Kilometer distances as in "km 82 T-Y" are measured from Tarapoto on the road to Yurimaguas. Otherwise place names are given in full.

TABLE 6

Genotypic data for *H. erato*

Site No.	$D^{Hy}D^{Hy}$ *Cr-* *Sd-*	$D^{Hy}D^{Hy}$ *Cr-* *sdsd*	*crcr* *Sd-*	*crcr* *sdsd*	$D^{Hy}d$ *Cr-* *Sd-*	$D^{Hy}d$ *Cr-* *sdsd*	*crcr* *Sd-*	*crcr* *sdsd*	*dd* *Cr-* *Sd-*	*dd* *Cr-* *sdsd*	*crcr* *Sd-*	*crcr* *sdsd*
1												18
2												4
3												2
4												12
5	5		1		1				1			
6	2				2		2			1		
7										1		210
8												3
9												8
9a												7
10												4
11												9
12												8
13												1
14												11[a]
17												2
18												2
19										1		131
20												1
21												1
22												38
23												16
24												2
25												15
26												22
27						2						26
28										1	1	15
29								3		1	2	19
30		1			1	1				3	3	22
31					2		1	2	2		2	17
32				1						2		
33	1				2		1	1	3	3	2	6
34	16	4	2		22	6	6	3	12	5	4	4
35	1				1		1		1	1		1
36	5	2			5		1		2			
37					2							
38								1		1		3
39							1		1	1	1	
40	4	1	1		8	2	1	1		3	1	
42	38	5	3		20	2	1		2	1	1	1
43	12		1		8	1	1	1	1			
44	90	5	2	3	37	2	2	1	1	1		
45	29	1			12			1	2			
46	8	2			1		1		1			
47	1				2							
48	86	3	5	1	26	2	2	1	1			
49	54	1	3		8			2				1
50	151	2	2		4	1						
51	1											
52	29		1									
53	20				1							
54	5	1										
55	20											

[a] In this site no hybrids between postman and rayed races of *erato* were found, although one interspecific hybrid with a parapatric sibling species, *Heliconius himera*, was captured.

TABLE 7

***H. melpomene* genotypic frequency data**

Site No.	D^R- N^N- *Yb-* *bb*	D^R- N^N- *Yb-* *B-*	*ybyb* *bb*	*ybyb* *B-*	N^BN^B *Yb-* *bb*	N^BN^B *Yb-* *B-*	*ybyb* *bb*	*ybyb* *B-*	*dd* N^N- *Yb-* *bb*	*dd* N^N- *Yb-* *B-*	*ybyb* *bb*	*ybyb* *B-*	N^BN^B *Yb-* *bb*	N^BN^B *Yb-* *B-*	*ybyb* *bb*	*ybyb* *B-*
1																20
3																11
4																4
6	4	1	1								1					2
7																26
8							3									16
9									1							8
9a									1							5
10																3
11																6
12																1
13																2
14																8
15																5
16																5
17																1
18																1
19							2									174
20																6
21																1
22							2									53
23																33
24																2
25																11
26																15
27																17
28							2									1
29		2					3		2							8
30	1	2	2		1		2		3							11
31		2					2		4							9
32	2	2					1				1					
33	10	12	2	1			5		8							10
34	5	7			1	1										3
35		1														
36	1	2					1									
39	8	7	1	5			1		2							3
40	11	4					1									
41		2			1		1									
42	22	14			2	1	1		2							1
43	39	18	2	3			4		2							2
44	17	12			1	1			1							1
45	7	3														
46	3	1														
47	4	1														
48	20	13	1			1										1
49	9	3							1							
50	3	3														
52	12	2														
53	8															
54	29															
55	4															