Hybrid Zones With Dobzhansky-Type Epistatic Selection

Sergey Gavrilets

Evolution, Vol. 51, No. 4. (Aug., 1997), pp. 1027-1035.

Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28199708%2951%3A4%3C1027%3AHZWDES%3E2.0.CO%3B2-H

Evolution is currently published by Society for the Study of Evolution.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ssevol.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Tue Apr 18 10:00:55 2006

EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 51 **August 1997** **No. 4**

Evolution, 51(4), 1997, pp. 1027–1035

HYBRID ZONES WITH DOBZHANSKY-TYPE EPISTATIC SELECTION

SERGEY GAVRILETS

Departments of Ecology and Evolutionary Biology and Mathematics,
University of Tennessee, Knoxville, Tennessee 37996-1300
E-mail: gavrila@ecology.tiem.utk.edu

Abstract.—Dobzhansky's model of epistatic selection assumes that viable genotypes form "clusters" in genotype space so that populations can evolve from one state to a reproductively isolated state following a "ridge" of well-fit genotypes without crossing any deep adaptive valleys. Recently, the importance of Dobzhansky-type models in evolutionary studies has been reemphasized by Gavrilets (1997a) and Gavrilets and Gravner (1997) who argue that the existence of "ridges" of well-fit genotypes connecting reproductively isolated genotypes is actually a general property of multidimensional adaptive landscapes. Using rigorous techniques and numerical simulations, I analyze clines in the frequencies of selected and neutral alleles maintained by a balance of migration and Dobzhansky-type epistatic selection acting on two diallelic loci. I show that Dobzhansky-type epistatic selection can build up a very strong barrier to neutral gene flow. I describe properties of clines that are indicative of Dobzhansky-type selection.

Key words.—Barrier to gene exchange, clines, epistasis, migration-selection balance.

Received August 9, 1996. Accepted April 9, 1997.

Analysis of hybrid zones provides insights into the nature of species, the strength and mode of natural selection, the genetic architecture of species differences, and the dynamics of the speciation process (Endler 1977; Barton and Hewitt 1981, 1985, 1989; Harrison and Rand 1989; Harrison 1990, 1993; Barton and Gale 1993). Many hybrid zones are thought to be formed following a secondary contact of different populations, and to be maintained by a balance between selection against hybrids and recombinant phenotypes and dispersal (Barton and Hewitt 1981, 1985, 1989; for alternative explanations see Endler 1977; Moore 1977; Rand and Harrison 1989). Several mathematical models intended to help to understand complex processes leading to the formation and fate of such hybrid zones have been proposed. Perhaps the most simple and basic model is that by Karlin and McGregor (1972), who considered two diploid populations connected by migration assuming that fitness is controlled by a single diallelic locus. In their model, individuals with genotypes **aa**, **aA**, and **AA** have relative fitnesses (viabilities) 1, $1 - s$, and 1, respectively ($s > 0$). Initially one allele is fixed in the first population whereas another allele is fixed in the second population. Given that migration rate m is small relative to the strength of selection ($m \leq m_c \approx s/6$, Karlin and McGregor 1972) two populations remain differentiated with alternative alleles close to fixation in each population.

Selection on some loci can produce clines in their allele frequencies that will represent a barrier to gene flow and will reduce the ability of (neutral) genes to penetrate from one population to another. The strength of the barrier to neutral gene flow in Karlin and McGregor's model was studied by Spirito et al. (1983) and Bengtsson (1985). Spirito et al. (1983) have characterized the strength of the barrier using the rate of convergence of the neutral allele frequency to an equilibrium value. To characterize the strength of the barrier Bengtsson (1985) defined a "gene flow factor" (GFF) as the probability of inclusion of an incoming neutral gene in a "native" genetic background. A more common measure is the strength of a genetic barrier between two populations, *b*, defined as the inverse of GFF (Barton and Bengtsson 1986). For the case of a neutral locus unlinked to locus **A**, Bengtsson's (1985) results give

$$b = \frac{1 + s}{1 - s}. \tag{1a}$$

For a neutral locus closely linked to locus **A** with recombination rate $R \ll 1$, Bengtsson's analysis produces

$$b = \frac{s}{R(1 - s)}. \tag{1b}$$

For the case of continuous habitat, the form of single-locus clines has been studied in Bazykin (1969), Barton (1979b), Christiansen et al. (1995), and Gavrilets (1997b). Moreover, Barton (1979a) and Piálek and Barton (1997) have considered gene flow across a hybrid zone. Several papers report theo-

© 1997 The Society for the Study of Evolution. All rights reserved.

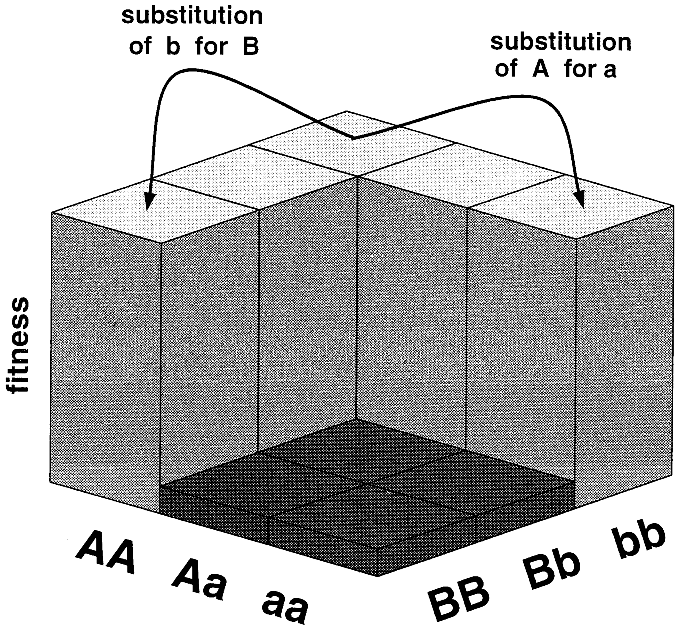

FIG. 1. Dobzhansky model (see text).

retical analyses of hybrid zones maintained by selection on multiple loci (Bazykin 1972a,b; Barton 1983, 1986; Bengtsson 1985; Barton and Bengtsson 1986; Barton and Gale 1993; Baird 1995; Gavrilets and Hastings 1996). The analyses differ mainly in the assumptions about the form of selection against hybrids and recombinant phenotypes.

Among different multilocus selection models a class of Dobzhansky-type epistatic models has become a subject of controversy. Dobzhansky's (1937) original model considers a two-locus, two-allele population initially monomorphic for a genotype, say **AAbb** (Fig. 1). This population is broken up into two geographically isolated parts. In one part, mutation and other factors cause substitution of **b** for **B** and a local race **AABB** is formed. In the other part, there is a substitution of **A** for **a**, giving rise to a local race **aabb**. It is assumed that there is no reproductive isolation among genotypes **AAbb**, **Aabb**, and **aabb** and among genotypes **AABB**, **AABb**, and **AAbb**, that is, all offspring of matings within these two groups are viable. In contrast, genotypes **AABB** and **aabb** are considered to be reproductively isolated in the sense that double heterozygote **AaBb** is inviable (or has very low fitness). In this scheme, strong reproductive isolation can be achieved, even though selection acting during the evolutionary divergence is weak or absent. Dobzhansky-type epistatic models (see Bengtsson and Christiansen 1983; Nei et al. 1983; Bengtsson 1985; Barton and Bengtsson 1986; Cabot at al. 1994; Wagner at al. 1994; Orr 1995; Orr and Orr 1996; Gavrilets and Hastings 1996) assume that viable genotypes form "clusters" in genotype space so that the population can move from one adaptive peak to another one separated by an adaptive valley following a "ridge" of well-fit genotypes without crossing any deep adaptive valleys. Populations diverge as a consequence of accumulation of different mutations (resulting from randomness of mutation and genetic drift) and reproductive isolation arises as a side effect of these accumulating differences between populations.

Bengtsson (1985) was the first to study the strength of the barrier to the neutral gene flow in a two-locus, Dobzhansky-type model. Using some heuristic approximations, Bengtsson found the gene flow factor for a neutral locus unlinked to the loci under selection. In the simple case of unlinked loci **A** and **B**, with perfectly viable genotypes along the "ridge" and all other genotypes having fitness $1 - s$, the strength of the barrier is given by equation (1a). This implies that if selection against F_1 hybrids is strong (i.e., if s is large), the barrier is strong. This conclusion was, however, revoked by Barton and Bengtsson (1986) who argued that Bengtsson's analysis described only the initial situation and that persistent migration between populations would quickly recreate the fitter intermediates along the ridge connecting the two adaptive peaks, destroying the barrier to gene exchange. Barton and Bengtsson (1986) presented a study of an epistatic model illustrating their point. On the other hand, a very short numerical study by Gavrilets and Hastings (1996) has indicated that Dobzhansky-type epistatic selection can build up a rather strong barrier to gene exchange.

Recently, it has been demonstrated that the existence of "ridges" of well-fit genotypes connecting reproductively isolated genotypes, which was postulated by Dobzhansky, is actually a general property of multidimensional adaptive landscapes (Gavrilets and Gravner 1997). A new metaphor of "holey" adaptive landscapes, which places a special emphasis on ridges of well-fit genotypes, has been put forward, and it has been argued that evolution and speciation on multidimensional adaptive landscapes proceeds according to the properties of underlying holey adaptive landscapes (Gavrilets 1997a; Gavrilets and Gravner 1997). This reinforces the importance of Dobzhansky-type models in evolutionary studies.

Here I study hybrid zones with Dobzhansky-type epistatic selection in a systematic way. Using rigorous techniques and numerical simulations I analyze clines in the frequencies of selected and neutral alleles maintained by a balance between migration and selection acting on two diallelic loci. I attempt to resolve whether Dobzhansky-type selection can produce a strong barrier to gene flow by examining the effects of migration rate on the strength of the barrier to gene exchange. I show that when selection against F_1 hybrids is weak, the fit intermediates can rise to high frequency, and isolation is lost. However, with strong selection, linkage disequilibrium can build up, and can keep the fit recombinants at low frequency, maintaining the barrier. I describe properties of clines in Dobzhansky-type models that might be useful in identifying the form of selection against hybrids and recombinants acting in natural populations.

THE MODEL

I consider a population with discrete generations under viability selection acting on two autosomal loci with alleles **A**,**a** and **B**,**b**, respectively. Let x_1, x_2, x_3, and x_4 be the frequencies of gametes **AB**, **Ab**, **aB**, and **ab**, respectively, and w_{jk} be the fitness (i.e., the probability of survival from zygote stage to maturity) of a zygote formed from gametes j and k. The gamete frequencies after selection and recombination in an isolated randomly mating population are described by the standard equations

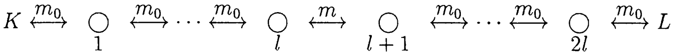

FIG. 2. Stepping-stone cline model (see text).

$$x_i' = \frac{w_i x_i \mp r w_{14} D}{\bar{w}}, \qquad i = 1, 2, 3, 4. \tag{2}$$

Here r is the rate of recombination, $D = x_1x_4 - x_2x_3$ is the standard linkage disequilibrium, $w_i = \Sigma_j\, w_{ij}x_j$ is the induced fitness of gamete i, and $\bar{w} = \Sigma_{ij}\, w_{ij}x_ix_j$ is the mean fitness of the population. In equation (2), the sign is minus for $i = 1$ and 4 and plus for $i = 2$ and 3.

Selection.—I assume that fitnesses of two-locus genotypes are given by matrix

$$\begin{array}{c|ccc} & \mathbf{BB} & \mathbf{Bb} & \mathbf{bb} \\ \hline \mathbf{AA} & 1 & 1-\beta & 1-\beta_1 \\ \mathbf{Aa} & 1-\alpha & 1-s & 1-\beta \\ \mathbf{aa} & 1-\alpha_1 & 1-\alpha & 1 \end{array}, \tag{3}$$

where α, α_1, β, β_1, and s are positive parameters that satisfy $\alpha \leq \alpha_1$, $\beta \leq \beta_1$, $\alpha \leq \beta$, s. In general, fitness matrix (3) describes an adaptive landscape with two equal peaks (at genotypes **AABB** and **aabb**) separated by an adaptive valley. Parameter s characterizes the degree of reproductive isolation between two populations at different adaptive peaks: the larger s is, the smaller is the fitness of F_1 hybrids, and the stronger is reproductive isolation. Dobzhansky-type epistatic models assume that strong reproductive isolation takes place simultaneously with the existence of a chain of almost perfectly viable genotypes connecting two "peaks." Within the framework of the fitness matrix (3), these assumptions will be incorporated by assuming that α and α_1 are much smaller than s, β and β_1, which are not small.

Fitness model (3) includes several previously studied models as partial cases. If $\alpha = \alpha_1 = 0$, one has a Dobzhansky-type model considered by Bengtsson and Christiansen (1983) and Bengtsson (1985). If $\alpha = \beta$, $\alpha_1 = \beta_1 = s$, one has a model with fitnesses depending on the proportion of foreign alleles, which was analyzed by Barton and Bengtsson (1986) and Barton and Gale (1993). Gavrilets and Hastings (1996) have considered a Dobzhansky-type model corresponding to the case with $s = 1$, $\alpha_1 = 2\alpha$, $\beta_1 = 2\beta$.

Spatial Structure.—I shall consider what happens after secondary contact of two populations at different adaptive peaks. I assume that two very large populations contact via a chain of smaller subpopulations. To model this situation, I use the stepping-stone cline model introduced by Feldman and Christiansen (1975). There are $2l$ subpopulations of equal size arrayed along a line connecting two main subpopulations with constant genotype frequencies (Fig. 2). The migration rate (i.e., the proportion of adults replaced by immigrants) between subpopulations of the same population is m_0, whereas the migration rate between "peripheral" subpopulations of different populations is m. I will assume that one of two main subpopulations, say that at left, consists of individuals with genotype **aabb**, whereas the second main subpopulation consists of individuals with genotype **AABB**. Initially all l subpopulations on the left have only **aabb** individuals whereas all l subpopulations on the right have only **AABB** individuals.

Let z_j be the frequency of a gamete in the jth subpopulation at the beginning of a generation, $j = 1, 2, \ldots 2l$. Gamete frequency $z_j{}'$ after selection and recombination is given by an equation in the form of (2). The gamete frequencies after adult migration are

$$z_l'' = (1 - m - m_0)z_l' + mz_{l+1}' + m_0z_{l-1}', \tag{4a}$$

$$z_{l+1}'' = (1 - m - m_0)z_{l+1}' + mz_l' + m_0z_{l+2}', \tag{4b}$$

$$z_j'' = (1 - 2m_0)z_j' + m_0(z_{j-1}' + z_{j+1}'), \qquad j \neq l,\ l + 1, \tag{4c}$$

with $j = 0$ and $j = 2l + 1$ corresponding to main subpopulations on the left and on the right, respectively. The case with $m_0 = 0$ and $l = 1$ describes direct contact of two populations.

ANALYTICAL RESULTS

I start by considering a simple case of $l = 1$, that is, when each population consists of a main subpopulation and a peripheral subpopulation with secondary contact taking place between the peripheral subpopulations. This case can also be considered as an approximation for a larger number of subpopulations on the time scale such that subpopulations away from the zone of contact are not yet affected by gene flow from other populations. Let x_1 through x_4 and X_1 through X_4 be the frequencies of the four gametes in the first and second peripheral subpopulations, respectively. Initially (i.e., before the contact between peripheral subpopulations) only gamete 4 (gamete **ab**) is present in the first subpopulation and only gamete 1 (gamete **AB**) is present in the second subpopulation. For the first peripheral subpopulation gamete 4 is "native," whereas gametes 1, 2, and 3 are "foreign." For the second peripheral subpopulation gamete 1 is "native," whereas gametes 2, 3, and 4 are "foreign." After individuals start migrating between peripheral subpopulations bringing genes foreign to each subpopulation's genetic background, an equilibrium cline in gamete frequencies will be formed. I consider different characteristics of this cline, including its ability to prevent neutral gene flow between populations. To gain some insight into these questions, I shall use a weak migration approximation and regular perturbations techniques, which were introduced into selection-migration studies by Svirezhev (1968) and Karlin and McGregor (1972) and have proved very useful since then.

Cline in Selected Loci

Cline Properties.—Let us assume that there is some migration between main subpopulations and their peripheral subpopulations ($m_0 \neq 0$) and that the migration rate between peripheral subpopulations, m, increases from zero to some small value. Exact values of equilibrium gamete frequencies cannot be found; therefore one is forced to use approximations. Standard regular perturbation techniques (e.g., Holmes 1995) produce the first-order approximations for the equilibrium gamete frequencies given in the appendix. These approximations show that any increase in selection (increase in α, β, s) or decrease in migration between populations m

always results in decreasing the frequencies of "foreign" gametes. This is what is expected from biological considerations. An increase in recombination rate r decreases the frequency of the "foreign" gamete complementary to the "native" gamete, but increases the frequencies of recombinant "foreign" gametes 2 and 3.

Let us assume that any deviation from an adaptive peak results in a substantial reduction in fitness (α, β, $s \gg m$). In this case, the equilibrium gamete frequencies are

$$x_1^* = X_4^* = \frac{m}{s + (1 - s)r}, \tag{5a}$$

$$x_2^* = X_2^* = \frac{(1 - s)r}{s + (1 - s)r} \frac{m}{\beta}, \tag{5b}$$

$$x_3^* = X_3^* = \frac{(1 - s)r}{s + (1 - s)r} \frac{m}{\alpha}, \tag{5c}$$

$$x_4^* = X_1^* = 1 - x_1^* - x_2^* - x_3^*. \tag{5d}$$

Frequencies of all foreign gametes are very small and have the same order (equal to the ratio of the migration rate and the strength of selection). In the case of Dobzhansky-type epistatic selection where strong selection against F_1 hybrids takes place simultaneously with existence of a chain of (almost) perfectly viable genotypes connecting two adaptive peaks (β, $s \gg m$, $m_0 \gg \alpha$) the equilibrium frequency of gamete **aB** becomes

$$x_3^* = X_3^* = \frac{(1 - s)r}{s + (1 - s)r} \frac{m}{m_0}. \tag{6}$$

The equilibrium frequencies of other foreign gametes are still given by equations (5a–d) and are small, whereas the frequency of gamete **aB** in both peripheral subpopulations is much bigger. Strictly speaking, the first-order approximations (5–6) are feasible as long as m remains much smaller than m_0. It is often the case, however, that first order analysis gives approximations reasonably good for broader parameter values. Equation (6) suggests that as the rate of migration between populations m increases to the level of that between main and peripheral subpopulations m_0, the frequency of gamete **aB** increases dramatically from a very low level (order m/s) to a significant level (order the ratio $r[1 - s]/[s + r(1 - s)]$). Numerical simulations have shown that with sufficiently large migration, the frequency of gamete **aB** reaches that of the "native" gametes and three genotypes **aabb**, **aaBb**, and **aaBB** become common in the first peripheral subpopulation, whereas three genotypes **AABB**, **AABb**, and **AAbb** become common in the second peripheral subpopulation.

Let p_1 and p_2 (P_1 and P_2) be the frequencies of allele **A** and **B** in the first (second) peripheral population. Since $p_1 = x_1 + x_2$ and $p_2 = x_3$, in the first subpopulation, the frequency of foreign allele **A** is very small but the frequency of foreign allele **B** is significant (and is approximately given by the right-hand side of eq. [6]). In a similar way, in the second peripheral subpopulation the frequency of the foreign allele **a** is significant. Using equations (5–6) one can show that equilibrium values of p_i and P_i satisfy

$$p_i^* = 1 - P_i^*, \qquad i = 1, 2, \tag{7a}$$

$$P_1^* - p_1^* = P_2^* - p_2^*, \tag{7b}$$

$$p_2^* > p_1^*, \qquad P_2^* > P_1^*. \tag{7c}$$

Equation (7a) is a consequence of the symmetry assumption incorporated in fitness matrix (3). Equation (7b) shows that changes between peripheral subpopulations in allele frequencies in the first and second loci are the same. In the case of continuous habitat, the width of a cline, w, is usually defined as the total change in gene frequency divided by the maximum gradient (Endler 1977). The discrete space analog of w in the model under consideration is one (the total change) divided by $P_i - p_i$ (the maximum change). Thus, equation (7b) implies that the width of the clines in allele frequencies in the first and second loci is the same. Equation (7c) tells us these clines are disjointed. Several other potentially important observations can be made. Populations at the zone of contact (peripheral subpopulations in the model under consideration) have genotypes that are absent (or are at very low frequencies) in main subpopulations and that are different from F_1 hybrids resulting from crosses between individuals from different main subpopulations. These subpopulations have large genetic variability in the loci under selection. However, this variability does not transform into large variability in fitness, and peripheral subpopulations have large mean fitness: $\bar{w} \approx 1 - 2m$. At the same time, reproductive isolation between two main subpopulations (as measured by the loss in F_1 hybrid fitness $w_{F1} = 1 - s$) can be very large.

Cline Stability.—Given that the rate of migration between main and peripheral subpopulations, m_0, is not zero, the cline in the gamete frequencies approximated by equations (5–6) is stable for any migration rate m and any adaptive landscape of form (3). This is not so if $m_0 = 0$, that is, if there is direct contact of two populations. To see this, let us consider the maximum eigenvalue λ_{max} of the stability matrix of the dynamic system (2–4) at equilibrium (5–6). In general, the equilibrium is locally stable if $-1 < \lambda_{max} < 1$. Let $\alpha \leq \beta$, s. Using standard regular perturbation methods, one finds that

$$\lambda_{max} \approx (1 - m_0)(1 - \alpha) + m\lambda_1, \tag{8}$$

where λ_1 depends on the parameters of the model. Increasing m should decrease the domain of stability and, thus, λ_1 should be positive. As $\alpha \to 0$, the domain of stability shrinks. If both $m_0 = 0$ and $\alpha = 0$, approximation (8) predicts that $\lambda_{max} > 1$ for any nonzero migration rate m. This means that in Dobzhansky-type models with $\alpha = 0$ and $m_0 = 0$ the nonuniform solution is unstable for any m-values. Bengtsson's (1985) analysis of the strength of genetic barrier in Dobzhansky-type model used these two assumptions, and thus Barton and Bengtsson's (1986) conclusion about collapse of the genetic barrier in Dobzhansky's model is correct. However, as I show below, Bengtsson's formulae are valid in a slightly different setup.

Neutral Gene Flow across The Hybrid Zone

Unless the loci under selection are identified, allele frequencies p_1 and p_2 cannot be measured. However, what can usually be found is a neutral marker locus (or loci) at which the populations away from the hybrid zone have clearly different frequencies. Information on clines in neutral loci observed in natural populations is abundant. To use this information in making biological conclusions, it is important to understand how neutral allele frequencies are expected to change across hybrid zones in different models.

Let us consider a third "neutral" locus M with alleles **m** and **M**. Let u and U be the frequency of allele **M** in the first and second peripheral subpopulations, respectively. I take the special order of loci to be ***MAB***. As before, let r be the recombination rate between the loci under selection A and B, and R be the recombination rate between M and A, and further assume that recombination occurs independently between the first, second, and third positions. Let the neutral allele M initially be absent in the first population, but fixed in the second population (i.e., initially $u = 0$, $U = 1$). After individuals begin migrating between the populations, the foreign neutral alleles start to appear in each subpopulation. In the case of a direct contact between populations (i.e., if $m_0 = 0$), the difference in the frequency of a neutral allele between populations is expected to gradually decay. In the stepping-stone cline model with $m_0 \neq 0$, a stable cline in the neutral allele frequency is expected to be formed. With no selection (i.e., if all parameters of the fitness matrix [3] are equal to zero), the equilibrium value of the neutral allele frequency in the first subpopulation is $u_{no} = m/(2m + m_0)$ reducing to $u_{no} = m/m_0$ if $m \ll m_0$ and to $u_{no} = 1/3$ if $m = m_0$. In the second subpopulation, $U_{no} = 1 - u_{no}$. Selection will result in deviation of equilibrium values of u and U from u_{no} and U_{no}, respectively.

Let y_1, y_2, y_3, and y_4 be the frequencies of the gametes **MAB**, **MaB**, **MBa**, and **Mab** in the first peripheral subpopulation (where initially gamete **mab** is fixed). These frequencies together with x_1, x_2, x_3, and x_4 values completely characterize the state of the first subpopulation. Obviously, $u = \Sigma\, y_i$. The appendix gives first-order approximations for y_i as well as relevant values for the second subpopulation found using regular perturbation techniques. Under the assumption that m and m_0 are small and about the same order, equilibrium frequencies of gametes **MAB**, **MAb**, and **MaB** are much smaller than the equilibrium frequency of gamete **Mab** (i.e., $y_1^*, y_2^*, y_3^* \ll y_4^*$). In this case, the frequency of a neutral allele **M** that is unlinked to the loci under selection ($R = 1/2$) in the first peripheral subpopulation is approximately given by

$$u^* \approx \frac{m}{m_0}\frac{1 - s}{1 + s + r(1 - s)} \times \left(1 - r + r\frac{1 - \alpha}{1 + \alpha} + r\frac{1 - \beta}{1 + \beta}\right). \tag{9a}$$

For Dobzhansky-type epistatic selection with $\alpha \ll 1$ and unlinked loci under selection ($r = 1/2$), this simplifies to

$$u^* \approx \frac{m}{m_0}\frac{(1 - s)(3 + \beta)}{(3 + s)(1 + \beta)}. \tag{9b}$$

For tightly linked selected loci ($r \ll 1/2$), (9a) reduces to

$$u^* \approx \frac{m}{m_0}\frac{1 - s}{1 + s}. \tag{9c}$$

The latter equation results also from (9b) if $\beta = s$. If $R = 1/2$, the frequency of **M** in the second peripheral subpopulation $U^* = 1 - u^*$. Approximations (9a–c) show that if selection against F_1 hybrids is strong (s is large), the foreign neutral allele does not penetrate the barrier (u^* is small) even it is unlinked to the loci under selection. Clines for the frequencies of different neutral alleles that are unlinked to the loci under selection coincides (and are characterized by eqs. [9a-c]).

On the other hand, if the neutral locus is closely linked to **A** (i.e., if $R \ll 1/2$),

$$u^* \approx \frac{m}{m_0}R\frac{1 - s}{s + r(1 - s)}\left(1 + r\frac{1 - \beta}{\beta}\right). \tag{10a}$$

If the loci under selection are unlinked (i.e., if $r = 1/2$),

$$u^* \approx \frac{m}{m_0}R\frac{(1 - s)(1 + \beta)}{(1 + s)\beta}. \tag{10b}$$

If all three loci are tightly linked (i.e., if $r \ll 1/2$), equation (10a) simplifies to

$$u^* \approx \frac{m}{m_0}R\frac{1 - s}{s}. \tag{10c}$$

If the neutral allele is linked to a locus under selection ($R < 1/2$), the equilibrium cline in the neutral allele frequency is not symmetric ($U^* \neq 1 - u^*$). Approximations (10a–c) show that if selection against F_1 hybrids is strong (s is large) or/and the neutral locus is closely linked to a selected locus ($R \ll 1/2$), the foreign neutral allele does not penetrate the barrier (u^* is small). The equilibrium frequency of a neutral allele depends on the degree of its linkage to a selected locus. Clines in the frequencies of selected alleles are disjointed. Thus, clines in the frequencies of different neutral alleles that are linked to the loci under selection are expected to be disjointed as well.

The results about the equilibrium values of the neutral allele frequency can be used to characterize the strength of the barrier to neutral gene flow across the hybrid zone. There are at least two approaches. First, one can introduce a measure characterizing the rate of convergence to the equilibrium values (e.g., based on eigenvalues; see Spirito et al. 1983). One can also introduce a measure based on the deviation of these equilibrium values from those expected with no selection. The latter approach is easier and is adapted here. I define the strength of the barrier to the flow of neutral genes from peripheral subpopulation 1 to peripheral subpopulation 2, b_+, and from peripheral subpopulation 2 to peripheral subpopulation 1, b_-, as

$$b_+ = (U^* - u^*)/u^*,\ b_- = (U^* - u^*)/(1 - U^*). \tag{11}$$

This definition is a discrete-space analog of that for populations in continuous habitat (Nagylaki 1976; Barton 1979a). The barrier strength, b, is the difference in the neutral allele frequency between two peripheral subpopulations divided by the difference in the neutral allele frequency between a peripheral subpopulation and its main subpopulation. It can be thought of as the number of subpopulations one should put between two peripheral subpopulations that would present an equivalent obstacle to the flow of a neutral allele (see Figs. 1c and 1d in Barton and Bengtsson 1986). With no selection,

$u^* = ⅓$, $U^* = ⅔$, and thus, $b = 1$. If the frequencies of foreign neutral alleles are small (i.e., if $u^* \ll 1$, $1 - U^* \ll 1$), then equation (11) simplifies to

$$b_+ \approx 1/u^*, \; b_- \approx 1/(1 - U^*). \tag{12}$$

Parameter configuration leading to equations (9b,c) have been considered by Bengtsson (1985). If the frequencies of foreign neutral alleles are small, the strength of the barrier to the gene flow is given by equation (12) and is different from that found by Bengtsson by a factor m/m_0.

Numerical Results

The methods used in the previous section do not allow one to evaluate the precision of analytical approximations when migration rates between (m) and within (m_0) populations are the same. In this section, I present numerical results illustrating properties of hybrid zones with Dobzhansky-type epistatic selection for the case $m_0 = m$.

Figure 3 shows how the equilibrium frequencies of selected alleles p_1 and p_2 and of the neutral allele u depend on the migration rate in three different selection models. Note that with extremely small migration all three allele frequencies should be about zero. With no selection, all allele frequencies should be equal to ⅓. Figures 3a,b correspond to the following fitness matrix

$$\begin{pmatrix} 1.0 & .99 & .98 \\ .99 & .98 & .99 \\ .98 & .99 & 1.0 \end{pmatrix}. \tag{13}$$

In this model any "path" from one adaptive peak to another involves the same reduction in fitness. This model will serve as a reference point for two Dobzhansky-type models considered below. Figures 3a,b show that the selected allele frequencies are equal and increase from zero (for small m) to a level close to that for the case with no selection (as m is about 0.01). Unless the migration rates are extremely small, the neutral allele frequency is close to ⅓. The barrier to neutral gene flow is practically absent. Close linkage between selected loci or between the neutral and a selected loci does not change this conclusion.

Figures 3c,d correspond to fitness matrix

$$\begin{pmatrix} 1.0 & .50 & .50 \\ .99 & .50 & .50 \\ .98 & .99 & 1.0 \end{pmatrix}. \tag{14}$$

In this model, there is a ridge of almost perfectly viable genotypes. Fitness of other genotypes is half of that of genotypes at adaptive peaks. Reproductive isolation is moderate (half of F_1 hybrids are inviable). Figures 3e,f correspond to the fitness matrix

$$\begin{pmatrix} 1.0 & .10 & .10 \\ .99 & .10 & .10 \\ .98 & .99 & 1.0 \end{pmatrix}. \tag{15}$$

Here again there is a ridge of almost perfectly viable genotypes. Fitness of other genotypes is $1/10$ of that of genotypes at adaptive peaks. Reproductive isolation is strong (only one of 10 F_1 hybrids is viable).

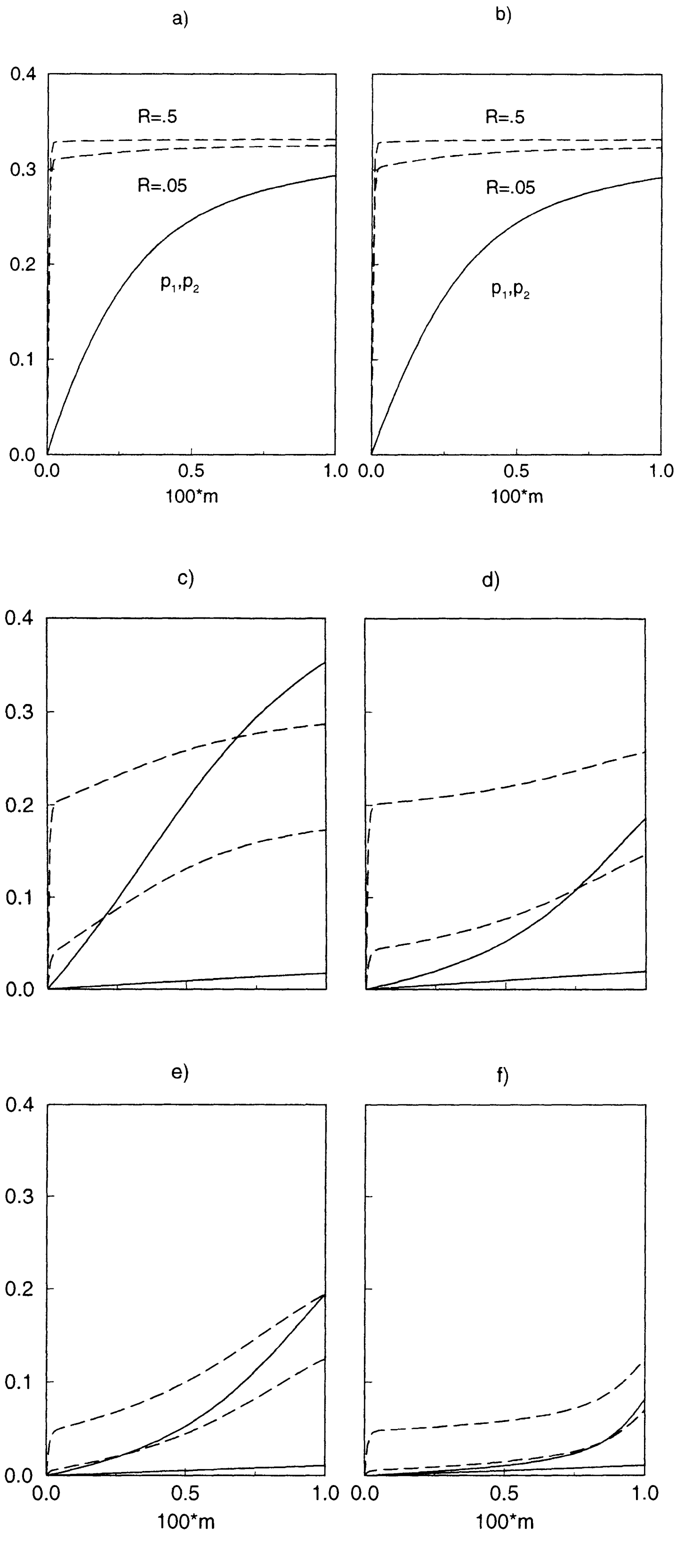

Fig. 3. Equilibrium allele frequencies p_1, p_2 (solid lines) and u (dashed lines) in the first peripheral subpopulation as functions of migration rate. In (a), (c), and (e) the loci under selection are unlinked ($r = ½$). In (b), (d), and (f) the loci under selection are closely linked ($r = 0.05$). Of the two solid lines, the upper gives p_1 and the lower gives p_2. In (a) and (b) the solid lines coinside. Of the two dashed lines, the upper gives u for $R = ½$ and the lower gives u for $R = 0.05$. See text for more details.

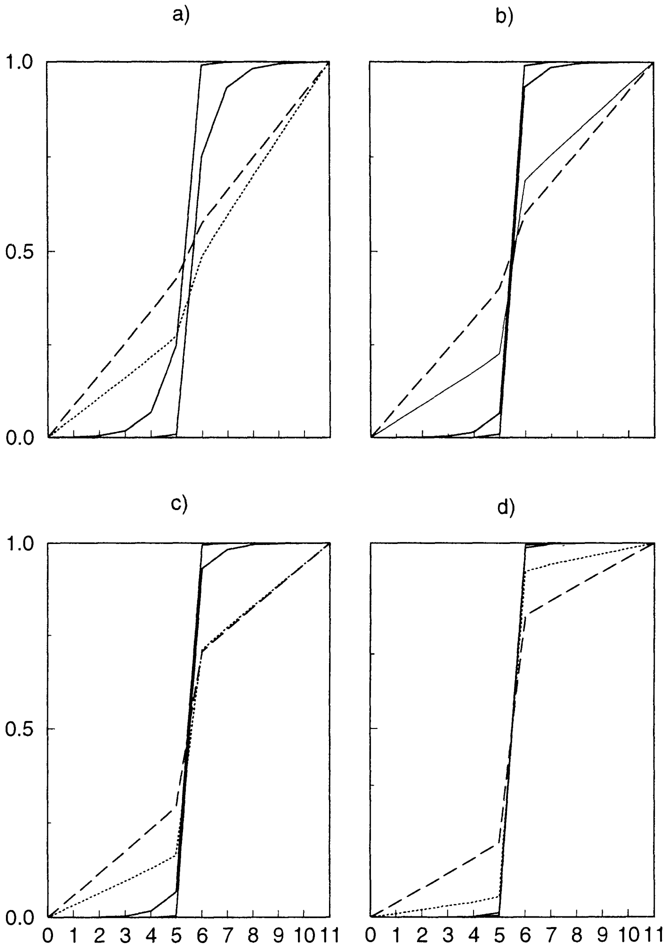

FIG. 4. Equilibrium clines in allele frequencies p_1, p_2, and u in a stepping-stone cline model with 10 subpopulations. Figures (a), (b) and (c), (d) corresponds to fitness matrices (14) and (15), respectively. In figures (a) and (c) the loci under selection are unlinked ($r = ½$). In figures (b) and (d) the loci under selection are closely linked ($r = 0.05$). Of the two solid lines, the upper gives p_1 and the lower gives p_2. The long-dashed line gives u for $R = ½$ and the short-dashed line gives u for $R = 0.05$. See text for more details.

In Figures 3c–f, the allele frequency p_1 is small throughout the considered range of m. The allele frequency p_2 increases from very low levels to a significant level. The neutral allele frequency u increases with m. The equilibrium allele frequencies are different between the three models. Thus, the barrier to gene exchange depends not only on fitnesses of intermediates along the ridge connecting the two adaptive peaks, which are the same in these models, but on fitnesses of F_1 hybrids as well (cf. Barton and Bengtsson 1986). Both linkage between neutral and selected loci and linkage between selected loci reduce equilibrium values of u. With strong linkage (Figs. 3c,d) or strong selection (Figs. 3d,e) u is small, and hence, the genetic barrier is strong. For example, in model (15) with $m = 0.005$ and unlinked loci, $u^* \approx .08$ (see Fig. 3e) giving $b \approx 12$. If all three loci are closely linked, $u^* \approx .02$ (see Fig. 3f) giving $b \approx 50$.

A different illustration of the same conclusions is given in Figure 4. This figure shows equilibrium values of p_1, p_2, and u in a system of 10 subpopulations. Here populations 0 and 11 correspond to main subpopulations with constant genotype frequencies. Note that with extremely small migration, all allele frequencies should be about zero in subpopulations 1–5 and about one in subpopulations 6–10. With no selection, all three allele frequencies should increase linearly from zero in the left main subpopulation to one in the right main subpopulation with increment $1/(2l + 1)$ (Feldman and Christiansen 1975, eq. 30), which is 1/11 with $l = 5$. Figures 4a,b and 4c,d corresponds to fitness matrices (14) and (15), respectively. The migration rate $m = 0.05$ was taken at the middle of the range considered in Figure 3. These Figures clearly illustrate that clines in selected allele frequencies can be disjointed.

DISCUSSION

This paper presents a systematic study of properties of hybrid zones maintained by a balance of migration and Dobzhansky-type epistatic selection. Previous theoretical studies of such hybrid zones have been limited in a number of ways. Bengtsson's (1985) pioneering analysis used heuristic approximations for the case of an unlinked neutral locus and a limited selection scheme. The cline analyzed by him was actually unstable (Barton and Bengtsson 1986). Barton and Bengtsson (1986) considered a specific Dobzhansky-type model assuming very weak selection and tight linkage. No attempt had been made in either paper to check the validity of approximations used. Gavrilets and Hastings's (1996) numerical study of the strong selection case was also very limited.

Here I used rigorous analysis to consider a more general selection scheme and a linked or unlinked neutral marker locus. I have checked analytical approximations against numerical simulations specifying the domain of applicability of the former. I have shown that the clines in the frequencies of selected alleles resulting from Dobzhansky-type epistatic selection represent a barrier to neutral gene flow. The strength of the barrier to gene exchange depends on the migration rate. With relatively large migration, the barrier to gene exchange depends mainly on fitnesses of intermediates along the ridge connecting the two adaptive peaks, and is not impressive. However, if migration rates are small, the barrier to gene exchange depends mainly on fitnesses of F_1 hybrids. If selection against F_1 hybrids is strong, linkage disequilibrium can build up and can keep the fit recombinants at low frequency, thus maintaining a very strong barrier. I believe these findings have resolved the controversy about the strength of the barrier to neutral gene flow in Dobzhansky-type models discussed in the introduction.

My results show that with strong Dobzhansky-type epistatic selection and low rates of migration, the following properties of hybrid zones should be expected. Reproductive isolation between allopatric populations on opposite sides of the hybrid zone (measured by the decrease in the fitness of F_1 hybrids) will increase with distance between these populations. F_1 hybrids between individuals from allopatric populations on opposite sides of the hybrid zone will have low fitness. These F_1 hybrids will have genotypes that differ from hybrid genotypes common in the center of the hybrid zone, which will have high fitness. In general, clines in the frequencies of neutral marker alleles linked to selected loci will be disjointed and unsymmetric. Concordant clines are expected for neutral alleles unlinked to selected loci. There have been many examples of hybrid zones with apparently discordant clines and apparently well-fitted recombinant genotypes present (e.g., Barton and Hewitt 1981, 1985; Harrison 1990; Bert and Arnold 1995; Nurnberger et al. 1995). A grasshopper hybrid zone studied by Virdee and Hewitt (1994) is especially interesting in this regard. Here crosses between the two pure taxa (*Chorthippus parallelus parallelus* and *Chorthippus parallelus erythropus*) result in sterile male offspring, whereas no such dysfunction has been detected in hybrid males collected through the center of the hybrid zone. Crosses have revealed noncoincident clines for dysfunction near the center of the hybrid zone.

The properties of hybrid zones maintained by balance of migration and Dobzhansky-type epistatic selection are different from those formed when adaptive peaks are isolated. Under this kind of epistasis any deviation from a "coadapted" combination of genes results in a (significant) reduction in fitness—there is no conditionally neutral substitutions. A major difference between the two types of hybrid zones should be in the distribution and fitnesses of genotypes in the center of the hybrid zone. If adaptive peaks are isolated, in the center of the hybrid zones beside the high-fitness parental forms, one should observe mainly low-fitness hybrids. Moreover, one expects concordant clines in neutral allele frequencies (Barton 1983, 1986; Nurnberger et al. 1995).

Detailed studies of the distribution and fitnesses of genotypes in the hybrid zone compared with those in crosses between distant populations as well as studies of the form and location of clines in different allele frequencies should be helpful in deducing the mechanisms responsible for hybrid zone formation and fate.

ACKNOWLEDGMENTS

I am grateful to N. Barton, M. Cruzan, R. Hallock, and D. Howard for helpful comments on the manuscript.

LITERATURE CITED

BAIRD, S. J. E. 1995. A simulation study of multilocus clines. Evolution 49:1038–1045.

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

———. 1979b. The dynamics of hybrid zones. Heredity 43:341–359.

———. 1983. Multilocus clines. Evolution 37:454–471.

———. 1986. The effect of linkage and density-dependent regulation on gene flow. Heredity 57:415–426.

BARTON, N. H., AND B. O. BENGTSSON. 1986. The barrier to genetic exchange between hybridizing populations. Heredity 56:357–376.

BARTON, N. H., AND K. S. GALE. 1993. Genetic analysis of hybrid zones. Pp. 13–45 *in* R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford Univ. Press, New York.

BARTON, N. H., AND G. M. HEWITT. 1981. Hybrid zones and speciation. Pp. 109–145 *in* W. R. Atchley and D. S. Woodruff, eds. Evolution and Speciation. Essays in honor of M. J. D. White. Cambridge Univ. Press, Cambridge.

———. 1985. Analysis of hybrid zones. Annu. Rev. Ecol. Syst. 16:113–148.

———. 1989. Adaptation, speciation and hybrid zones. Nature 341:497–503.

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

———. 1972a. The disadvantage of heterozygotes in a population with a continuous area. Genetika 8:162–167.

———. 1972b. The disadvantage of heterozygotes in a system of two adjacent populations. Genetika 8:155–161.

BENGTSSON, B. O. 1985. The flow of genes through a genetic barrier. Pp. 31–42 *in* J. J. Greenwood P. H. Harvey, and M. Slatkin, eds. Evolution essays in honor of John Maynard Smith. Cambridge Univ. Press, Cambridge.

BENGTSSON B. O., AND F. B. CHRISTIANSEN. 1983. A two-locus mutation selection model and some of its evolutionary implications. Theor. Popul. Biol. 24:59–77.

BERT, T. M., AND W. S. ARNOLD. 1995. An empirical test of predictions of two competing models for the maintenance and fate of hybrid zones: both models are supported in a hard-clam hybrid zone. Evolution 49:276–289.

CABOT, E. L., A. W. DAVIS, N. A. JOHNSON, AND C.-I. WU. 1994. Genetics of reproductive isolation in the *Drosophila simulans* clade: complex epistasis underlying hybrid male sterility. Genetics 137:175–189.

CHRISTIANSEN, F. B., V. ANDREASEN, AND E. T. POULSEN. 1995. Genotypic proportions in hybrid zones. J. Math. Biol. 33:225–249.

DOBZHANSKY, TH. 1937. Genetics and the origin of species. Columbia Univ. Press, New York.

ENDLER, J. A. 1977. Geographic variation, speciation, and clines. Princeton Univ. Press, Princeton, NJ.

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

GAVRILETS, S. 1997a. Evolution and speciation on holey adaptive landscapes. Trends Ecol. Evol. *In press.*

———. 1997b. Single locus clines. Evolution 51:979–983.

GAVRILETS, S., AND J. GRAVNER. 1997. Percolation on the fitness hypercube and the evolution of reproductive isolation. J. Theoret. Biol. 184:51–64.

GAVRILETS, S., AND A. HASTINGS. 1996. Founder effect speciation: a theoretical reassessment. Am. Nat. 147:466–491.

HARRISON, R. G. 1990. Hybrid zones: windows on evolutionary process. Pp. 69–128 *in* D. Futuyma and J. Antonovics, eds. Oxford surveys in evolutionary biology. Vol. 7. Oxford Univ. Press, New York.

———. 1993. Hybrid zones and the evolutionary process. Oxford Univ. Press, New York.

HARRISON, R. G., AND D. M. RAND. 1989. Mosaic hybrid zones and the nature of species boundaries. Pp. 111–133 *in* D. Otte and J. Endler, eds. Speciation and its consequences. Sinauer Associates, Sunderland, MA.

HOLMES, M. H. 1995. Introduction to perturbation methods. Springer-Verlag, New York.

KARLIN S., AND J. MCGREGOR. 1972. Application of method of small parameters in multi-niche population genetics models. Theor. Popul. Biol. 3:180–209.

———. 1974. Towards a theory of the evolution of modifier genes. Theor. Popul. Biol. 5:59–103.

MOORE, W. S. 1977. An evaluation of narrow hybrid zones in vertebrates. Quart. Rev. Biol. 52:263–277.

NAGYLAKI, T. 1976. Clines with variable migration. Genetics 83: 867–886.

NEI, M., T. MARUYAMA, AND C-I. WU. 1983. Models of evolution of reproductive isolation. Genetics 103:557–579.

NURNBERGER, B., N. BARTON, C. MACCALLUM, J. GILCHRIST, AND M. APPLEBY. 1995. Natural selection on quantitative traits in the *Bombina* hybrid zone. Evolution 49:1224–1238.

ORR, H. A. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:1803–1813.

ORR, H. A., AND L. H. ORR. 1996. Waiting for speciation—the effect of population subdivision on the time to speciation. Evolution 50:1742–1749.

PIÁLEK, J., AND N. H. BARTON. 1997. The spread of an advantageous allele across a barrier: the effect of random drift and selection against heterozygotes. Genetics 145:493–504.

RAND, D. M., AND R. G. HARRISON. 1989. Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. Evolution 43:432–449.

SPIRITO, F., C. ROSSI, AND M. RIZZONI. 1983. Reduction of gene flow due to partial sterility of heterozygotes for a chromosome mutation. I. Studies on a "neutral" gene not linked to the chromosome mutation in a two population model. Evolution 37:785–797.

SVIREZHEV, Y. M. 1968. The systems of weakly connected populations. Stud. Biophys. 10:25–30.

VIRDEE, S. R., AND G. M. HEWITT. 1994. Clines for hybrid dysfunction in a grasshopper hybrid zone. Evolution 48:392–407.

WAGNER, A., G. P. WAGNER, AND P. SIMILION. 1994. Epistasis can facilitate the evolution of reproductive isolation by peak shifts: a two-locus two-allele model. Genetics 138:533–545.

Corresponding Editor: D. Howard

APPENDIX

With two alleles at each of the three loci, there are eight different gametes. The vector of eight gamete frequencies characterizes the state of a subpopulation. Equations describing local dynamics for two selected and one neutral locus are given in Karlin and McGregor (1974). The dynamics of the system with two peripheral subpopulations are described by 14 difference equations. With no migration between peripheral subpopulations (i.e., if $m = 0$), the dynamical system describing the model under consideration has a locally stable equilibrium with gametes **mab** and **MAB** fixed in the first and second peripheral subpopulation, respectively. The idea of regular perturbation techniques (e.g., Holmes 1995) is to approximate equilibrium values, eigenvalues, and other characteristics of the full system (with $m \neq 0$) as small deviations from those of the simplified system (with $m = 0$). Using these techniques gives the following first-order approximations for the frequencies of gametes **AB**, **Ab**, **aB**, and **ab** in the first peripheral subpopulation

$$x_1^* = \frac{m}{1 - (1 - m_0)(1 - s)(1 - r)}, \tag{A1a}$$

$$x_2^* = \frac{(1 - m_0)(1 - s)}{m_0 + \beta(1 - m_0)} r x_1^*, \tag{A1b}$$

$$x_3^* = \frac{(1 - m_0)(1 - s)}{m_0 + \alpha(1 - m_0)} r x_1^*, \tag{A1c}$$

$$x_4^* = 1 - x_1^* - x_2^* - x_3^*. \tag{A1d}$$

The frequencies of these gametes in the second subpopulation are $X_1^* = x_4^*$, $X_2^* = x_3^*$, $X_3^* = x_2^*$, $X_4^* = x_1^*$. Assuming $m_0 \ll \alpha, \beta$ one gets equations (5a–d). Assuming $\alpha \ll m_0 \ll \beta$, one gets equation (6).

One finds the following first-order approximations for the frequencies of gametes **MAB**, **MAb**, **MaB**, and **Mab** in the first peripheral subpopulation

$$y_1^* = \frac{m}{1 - (1 - m_0)(1 - R)(1 - r)(1 - s)}, \tag{A2a}$$

$$y_2^* = \frac{(1 - m_0)(1 - R)r(1 - s)}{1 - (1 - m_0)(1 - R)(1 - \beta)} y_1^*, \tag{A2b}$$

$$y_3^* = \frac{(1 - m_0)Rr(1 - s)}{1 - (1 - m_0)[(1 - R)(1 - \alpha) - (1 - 2R)r(1 - s)]} y_1^*, \tag{A2c}$$

$$\begin{aligned} m_0 y_4^* = {} & (1 - m_0)R(1 - r)(1 - s)y_1^* + (1 - m_0)R(1 - \beta)y_2^* \\ & + (1 - m_0)[R(1 - \alpha) + (1 - 2R)r(1 - s)]y_3^*. \end{aligned} \tag{A2d}$$

If $m_0 \ll 1$ and $R = 1/2$, the latter system of equations simplifies to

$$y_1^* = \frac{2m}{2 - (1 - r)(1 - s)}, \tag{A3a}$$

$$y_2^* = \frac{r(1 - s)}{1 + \beta} \frac{m}{2 - (1 - r)(1 - s)}, \tag{A3b}$$

$$y_3^* = \frac{r(1 - s)}{1 + \alpha} \frac{m}{2 - (1 - r)(1 - s)}, \tag{A3c}$$

$$y_4^* = \frac{m}{m_0} \frac{1 - s}{2 - (1 - r)(1 - s)} \left(1 - r + r\frac{1 - \alpha}{1 + \alpha} + r\frac{1 - \beta}{1 + \beta}\right). \tag{A3d}$$

Let Y_1, Y_2, Y_3, and Y_4 be the frequencies of gametes **mAB**, **mAb**, **maB**, and **mab** in the second peripheral subpopulation. Using regular perturbation techniques, one finds the following first-order approximations for the equilibrium values

$$Y_1^* = y_1^*, \tag{A4a}$$

$$Y_2^* = \frac{(1 - m_0)Rr(1 - s)}{1 - (1 - m_0)[(1 - r)(1 - \beta) + (1 - 2r)R(1 - s)]} y_1^*, \tag{A4b}$$

$$Y_3^* = \frac{(1 - m_0)R(1 - r)(1 - s)}{1 - (1 - m_0)(1 - r)(1 - \alpha)} y_1^*, \tag{A4c}$$

$$\begin{aligned} m_0 Y_4^* = {} & (1 - m_0)(1 - R)r(1 - s)Y_1^* \\ & + [r(1 - \beta) + (1 - 2r)R(1 - s)]Y_2^* + r(1 - \alpha)Y_3^*. \end{aligned} \tag{A4d}$$

Figure A1 demonstrates that the first-order approximations (A1–A4) predict the numerically found equilibrium allele frequencies reasonably well.

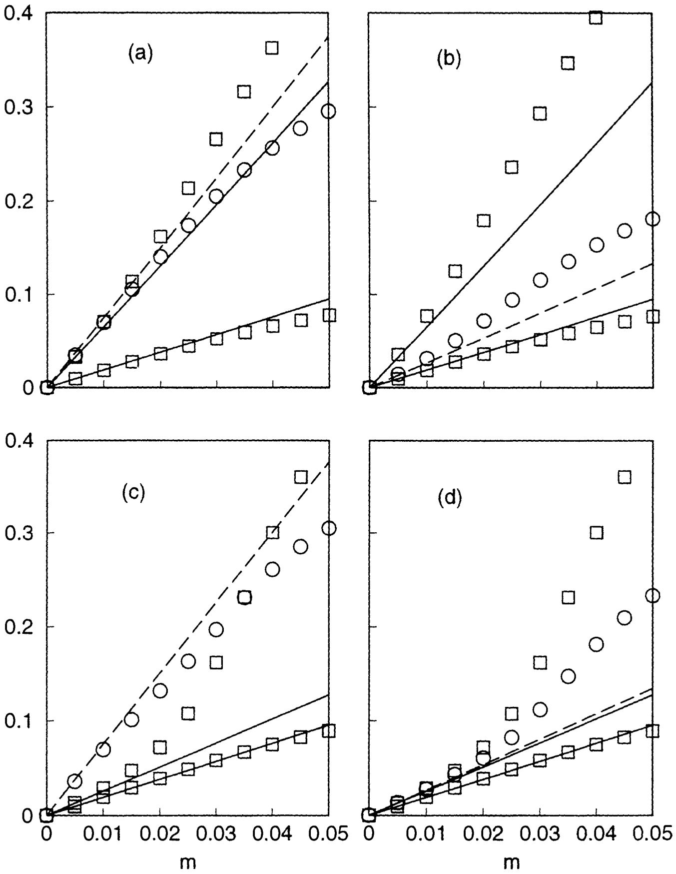

FIG. A1. Comparison of the equilibrium allele frequencies p_1, p_2, and u as functions of migration rate predicted by the first-order analysis and observed in numerical simulations. Of the two solid lines, the upper gives analytical approximation for p_1 and the lower gives analytical approximation for p_2. The long-dashed line gives analytical approximation for u. Squares and circles denote numerically found equilibrium values of the frequencies of selected and neutral alleles, respectively. See text for more details.