Genet. Res., Camb. (2000), **76**, *pp.* 179–185. *With* 1 *figure.* Printed in the United Kingdom © 2000 Cambridge University Press

The strength of the selection barrier between populations

KONSTANTIN V. PYLKOV[1], LEV A. ZHIVOTOVSKY[1]
AND FREDDY BUGGE CHRISTIANSEN[2]*
[1] *N. I. Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, Russia*
[2] *Department of Biology, University of Aarhus, Ny Munkegade, Building 540, DK-8000 Aarhus, Denmark*

(*Received 16 September 1999 and in revised form 13 March 2000*)

Summary

Genetic differences among populations exposed to selection form barriers against genetic exchange by mortality among hybrids. The strength of such a selection barrier, with which one (recipient) population reacts against immigration from another (donor) population, may be measured as the cumulative mean fitness of hybrids and their descendants relative to the fitness of the recipient population. Previous work analysed a case of weak selection with pairwise epistatic interactions by assuming small genetic distance between two populations in contact. The present study allows large genetic difference between the donor and recipient populations and considers weak multilocus selection with arbitrary epistatic interactions between two or more linked loci. An approximate analytical expression for the barrier strength is obtained as an expansion in which the strength of selection plays the role of a small parameter. It is shown that allele frequencies and gametic linkage disequilibria contribute in different ways to the strength of the selection barrier.

1. Introduction

Biological species are separated by various mechanisms of reproductive isolation that decrease the exchange of genetic material among them, and two principal types of isolation mechanisms, prezygotic and postzygotic, are usually distinguished. A prezygotic barrier is the inability of individuals to perform interspecific matings due to physiological and behavioural differences. A postzygotic mechanism of isolation is a lower viability or fertility of hybrids or hybrid descendants due to genetic difference between the species.

When two genetically distinct populations that are in contact produce hybrids with reduced fitness, they are less susceptible to mutual introgression of genetic material. The resulting genetic barrier therefore amounts to decreased effective migration between the populations compared with the exchange of breeding individuals. The added genetic isolation of the barrier promotes genetic divergence between populations, and this in turn contributes to an amplification of the barrier. These aspects in the evolution of a genetic barrier require quantitative study in order to understand their role in the speciation process, and thus in the progressive deepening of the postzygotic barrier as the species diverge.

The theory of genetic barriers is immediately applicable to the exploration of the impact of genetically modified organisms that are introduced into natural habitats where native populations may be susceptible to introgression. The specific problems of the evolutionary fate of the genetic material of modified individuals are not different from the problems of how genetic differences influence the survival of genetic material from one population in the environment of another divergent population.

Bengtsson (1985) quantified the strength of a genetic barrier by describing the fate of a neutral gene carried into a population by recurrent immigration, and this was later extended to models of hybrid zones (Barton & Bengtsson, 1986). Even though a neutral gene is irrelevant to the selection pressure against hybrids, it is nevertheless indirectly selected because of its association with the factors involved in the determination of the fitnesses of hybrids and their

* Corresponding author.

descendants. The strength of a genetic barrier was defined as its ability to protect the population against immigration of the neutral allele, that is, as the ratio of the migration rate and the effective migration rate describing the actual influx of the marker gene.

Most models involving hybrids assume reduced hybrid fitness (Bazykin, 1969; Barton, 1979; Barton & Hewitt, 1985; Christiansen *et al*., 1995). A different approach has been developed by Zhivotovsky & Christiansen (1995), in which the fitness of hybrids is determined by their multilocus genotypes and the selection regime in the recipient population. The only assumption is that the recipient population is at a stable genetic equilibrium. The reduced fitness of hybrids is therefore not an *a priori* assumption, and increased fitness may indeed result, because the mean fitness at equilibrium is unlikely to be at a local maximum. The barrier is quantified by considering a small group of individuals that immigrate into a genetically distinct recipient population. The immigrants mate at random with the residents, producing hybrids which in turn mate with the recipients, and so on. The immigration amounts to a small perturbation of the recipient population away from its stable equilibrium, and it therefore returns to the original equilibrium. Denote the mean fitness of the hybrid subpopulation at generation t by $\bar{w}_t$, and let $\bar{w}^*$ be the equilibrium mean fitness in the recipient population. Zhivotovsky & Christiansen (1995) describe the fate of immigrants and their descendants by introducing the *selection barrier* which describes the ultimate survival of hybrids relative to the survival of a similar number of residents:

$$\prod_{t=t_0}^{\infty} \frac{\bar{w}_t}{\bar{w}^*}. \tag{1}$$

Assuming equal fertility of each genotype, this expression gives the ultimate factor of change in the number of hybrids relative to that of the same number of recipients. The strength of the selection barrier is quantified by the rate of change:

$$B = -\ln\left(\prod_{t=t_0}^{\infty} \frac{\bar{w}_t}{\bar{w}^*}\right). \tag{2}$$

If $B > 0$, the fitness of hybrids and their descendants averaged over time is lower than the average fitness in the recipient population. Thus, the genetic impact of the immigrants will be lower than suggested by their number. If $B < 0$, the time-average fitness of hybrids and their descendants is higher than the average fitness in the recipient population. The impact of immigration will therefore be higher than expected, but the recipient population will nevertheless return to its original stable equilibrium.

Zhivotovsky & Christiansen's (1995) measure B of the strength of the selection barrier corresponds closely to that suggested by Bengtsson (1985) if the diagnostic neutral locus is *unlinked* to any of the loci involved in creating the selection barrier (Christiansen, 1999). Bengtsson's (1985) and Barton & Bengtsson's (1986) measure describes more accurately the rate of immigration of a neutral allele, because it allows linkage between the neutral locus and loci determining the fitness of the individuals. Zhivotovsky & Christiansen's (1995) measure is easier to evaluate in a theoretical analysis, and it is therefore used to provide a general impression of the barrier between two populations. Thus, the value of B reveals the susceptibility of the recipient population $\mathscr{R}$ to an introgression of the different genotypes carried from the donor population $\mathscr{D}$.

Zhivotovsky & Christiansen (1995) calculated the strength of the selection barrier between genetically close populations for weak selection with only some pairwise epistatic interactions allowed. We extend their results to a general weak multilocus selection model with arbitrary genetic distance between the donor population and the recipient population.

2. The model

We consider a randomly mating recipient population $\mathscr{R}$ with diploid individuals that may differ at n diallelic loci subject to viability selection. Selection is assumed to be weak, in that changes in the genetic composition of the population due to selection are slow compared with the changes mediated by recombination among the loci. The recipient and donor populations may therefore be assumed to be at quasi linkage equilibrium prior to the event of immigration (Kimura, 1965; Nagylaki, 1992; Nagylaki *et al*., 1999). Denoting the two possible alleles 0 and 1 at each locus, we describe the state of locus i by an indicator variable l_i, $i = 1, 2, \ldots, n$. This enables us to represent a gamete as $l = (l_1, l_2, \ldots, l_n)$ and the genotype of an individual as $l'l''$. The fitness of genotype $l'l''$ is $w_{l'l''} = \mu + \epsilon\nu_{l'l''}$, where ϵ is assumed small. Population $\mathscr{R}$ is assumed to be at an equilibrium that is stable under the joint action of selection and recombination.

A small group of immigrants is introduced from a genetically different donor population $\mathscr{D}$ into the recipient population $\mathscr{R}$ after selection, and these immigrants mate at random with residents and produce hybrids. To investigate the fate of the descendants of the immigrants, we consider the group of hybrids, $\mathscr{H}$, formed by the individuals with immigrant ancestors (Zhivotovsky & Christiansen, 1995). Because the immigrant group is assumed to be small and because random mating prevails, we shall not consider matings among the hybrid individuals. Immigration takes place at the initial generation 0, so the descendants of immigrants in generation 1 are produced from matings of immigrants and residents.

https://doi.org/10.1017/S001667230000464X Published online by Cambridge University Press

Thus, we describe the hybrid population in the subsequent generations by the recursion $\mathscr{H}_{t+1} = \mathscr{H}_t \times \mathscr{R}$ and $\mathscr{H}_0 = \mathscr{D}$. At the loci subject to selection, the genotypic composition of the hybrid subpopulation $\mathscr{H}$ will converge to that of the undisturbed recipient population $\mathscr{R}$ in the course of time.

The gene frequencies, linkage disequilibria and the mean fitness in population $\mathscr{H}_t$ are $p_i^{(t)}$, $D_K^{(t)}$ and $\bar{w}_t$ respectively, and when convenient we will suppress the reference to the generation number t. The linkage disequilibrium D_K refers to that among the loci in the subset K of the set of loci $N = \{1, 2, \ldots, n\}$ (Slatkin, 1972; Christiansen, 1999). The values of the variables at equilibrium in the recipient population $\mathscr{R}$ are p_i^*, D_K^* and $\bar{w}^*$ and in the donor population $\mathscr{D}$ they are p_i^{**} and D_K^{**}. We describe the dynamics in $\mathscr{H}_t$ by the deviation of the variables from their values at equilibrium in the recipient population $\mathscr{R}$:

$$\begin{aligned} \Delta p_i &= p_i - p_i^* \\ \Delta D_K &= D_K - D_K^* \\ \Delta \bar{w} &= \bar{w} - \bar{w}^*. \end{aligned} \tag{3}$$

The initial values of these deviations are therefore $\Delta p_i^0 = p_i^{**} - p_i^*$ and $\Delta D_K^0 = D_K^{**} - D_K^*$.

After t generations of weak selection, the mean fitness among hybrids $\mathscr{H}_t$ is $\bar{w}_t = \bar{w}^* + \Delta_t \bar{w}$, and as $\Delta_t \bar{w}$ is small, at most on the order of ϵ, we get

$$\log \frac{\bar{w}_t}{\bar{w}^*} = \log\left(1 + \frac{\Delta_t \bar{w}}{\bar{w}^*}\right) \approx \frac{\Delta_t \bar{w}}{\bar{w}^*} - \frac{1}{2}\left(\frac{\Delta_t \bar{w}}{\bar{w}^*}\right)^2.$$

Hence, for weak selection the value of selection barrier (2) is approximately

$$B \approx -\sum_{t=1}^{\infty} \frac{\Delta_t \bar{w}}{\bar{w}^*}, \tag{4}$$

where the error of the approximation is on the order of magnitude of ϵ^2.

The difference in mean fitness between the hybrid subpopulation $\mathscr{H}$ and the recipient population $\mathscr{R}$ is

$$\Delta w = \tfrac{1}{2} \sum_{i \in N} \Delta p_i \frac{\partial \bar{w}^*}{\partial p_i^*} + \tfrac{1}{2} \sum_{\substack{K \subseteq N \\ |K| > 1}} \delta_K \frac{\partial \bar{w}^*}{\partial D_K^*} + \mathscr{O}(\epsilon^2), \tag{5}$$

where the second sum is taken over all subsets of loci containing at least two elements. This is shown by using equations (2), (4) and (8) of Zhivotovsky & Pylkov (1998) and the assumption that population $\mathscr{R}$ is at equilibrium. The variables used in the linkage disequilibrium term of (5) are given by

$$\begin{aligned} \delta_{ij} &= \Delta D_{ij} + \Delta p_i \Delta p_j, \\ \delta_{ijk} &= \Delta D_{ijk} + \Delta D_{ij} \Delta p_k + \Delta D_{ik} \Delta p_j + \Delta D_{jk} \Delta p_i \\ &\quad + \Delta p_i \Delta p_j \Delta p_k, \\ &\vdots \\ \delta_K &= \sum_{\substack{V \subseteq K \\ |V| > 1}} \Delta D_V \prod_{i \in K \setminus V} \Delta p_i + \prod_{i \in K} \Delta p_i. \end{aligned} \tag{6}$$

These are combined measures of differences in linkage disequilibria, and through the back-cross series $\mathscr{H}' = \mathscr{H} \times \mathscr{R}$, the dynamics of these variables satisfy the simple recursion

$$\delta_K' = \tfrac{1}{2}(1 - r_K)\, \delta_K, \tag{7}$$

where r_K is the probability of at least one recombination event among the loci in K. This equation can be derived from equation (25) of Zhivotovsky & Pylkov (1998). The analysis of genetic changes in population $\mathscr{R}$ generalizes that of Zhivotovsky & Gavrilets (1992) and is closely related to those applied by Barton & Turelli (1991) and Turelli & Barton (1994).

Using (4), (5) and (7) we obtain the strength of the selection barrier as

$$B \approx \frac{1}{2\bar{w}^*} \left(\sum_{i \in M} \left| \left(\frac{\partial \bar{w}^*}{\partial p_i^*} \right) \Delta p_i^0 \right| - \sum_{\substack{K \subseteq N \\ |K| > 1}} \frac{1 - r_K}{1 + r_K} \left(\frac{\partial \bar{w}^*}{\partial D_K^*} \right) \delta_K^0 \right), \tag{8}$$

where M is the set of loci that are monomorphic in population $\mathscr{R}$. If $|\Delta p_i^0| \gg \epsilon$, B is on the order of ϵ, and the error in the approximation is of the order ϵ^2.

Thus, the barrier is produced by the selection gradient in the recipient population multiplied by the corresponding measures of differences between the donor and recipient populations. The one-locus selection gradient provides a positive contribution to the barrier for loci that are monomorphic in the recipient population. The sign of the contribution due to the epistatic interactions among the loci is undetermined; the selection gradient depends only on the genotypic fitnesses in the recipient population, whereas the measures of linkage disequilibrium depend on the difference between the recipient population and the original immigrants from the donor population.

(i) *The additive-by-additive epistasis model*

The analysis of Zhivotovsky & Christiansen (1995) was based on a simplified model which allows only additive and dominance effects within loci, and additive-by-additive epistasis among them according to the genotypic fitness model by Zhivotovsky & Gavrilets (1992):

$$\begin{aligned} w_{l'l''} &= \mu + \sum_{i=1}^{n} (a_i (l_i' + l_i'') + 2 b_i l_i' l_i'') \\ &\quad + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} c_{ij} (l_i' + l_i'')(l_j' + l_j''). \end{aligned} \tag{9}$$

The fitness of individuals carrying only 0 alleles at every locus is μ, and the coefficient a_i measures the additive contribution of allele 1 at locus i and b_i the dominance contribution. The coefficient $c_{ij}, i \neq j$

https://doi.org/10.1017/S001667230000464X Published online by Cambridge University Press

Table 1. *The three donor populations used in iterations*

Donor population $\mathscr{D}$		All 1-allele frequencies	All pairwise linkage disequilibria
A	Two complementary gametes $(0,0,0,0,0):\frac{1}{2}$ and $(1,1,1,1,1):\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$
B	All gametes in equal frequencies $\frac{1}{32}$	$\frac{1}{2}$	0
C	Single gamete $(1,1,1,1,1)$	1	0

measures the additive-by-additive epistatic interaction of the effects of the 1-alleles at loci i and j. The coefficients a_i, b_i and c_{ij} are at most on the order of magnitude of ϵ.

In the additive-by-additive epistasis model the selection barrier (8) becomes

$$B \approx \frac{1}{\bar{w}^*}\left(\sum_{i\in M} |F_i^*||\Delta p_i^o| - \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{1-r_{ij}}{1+r_{ij}} \times c_{ij}(\Delta D_{ij}^o + \Delta p_i^o \Delta p_j^o) \right), \quad (10)$$

where

$$F_i^* = a_i + 2b_i p_i^* + \sum_{\substack{j=1\\ j\neq i}}^{n} 4c_{ij} p_j^*.$$

The simplification of the first term is founded on the property that $F_i^* \Delta p_i^o \leqslant 0$ for all $i \in M$ and $F_i^* = 0$ otherwise. The order of magnitude of F_i^* is ϵ, and if the difference in gene frequencies between $\mathscr{R}$ and $\mathscr{D}$ is large ($|\Delta p_i^o| \gg \epsilon$) then B becomes of the order ϵ. Equation (10) differs from the result of Zhivotovsky & Christiansen (1995) by the term $\Delta p_i^o \Delta p_j^o$, which becomes negligible if the difference in the allele frequencies is indeed on the order of ϵ as they assumed.

The loci monomorphic in population $\mathscr{R}$ always contribute to the first and positive term in (10), and this term is independent of recombination. The second term depends on recombination and epistatic interactions. It consists of two parts, namely a term with $\Delta p_i^o \Delta p_j^o$ which is proportional to the covariance in gene frequencies between the two populations, and a term with ΔD_{ij}^o which is the difference in linkage disequilibria between $\mathscr{R}$ and $\mathscr{D}$.

The covariance in gene frequencies between the two populations emerges because population $\mathscr{H}$ is a result of mixing populations $\mathscr{R}$ and $\mathscr{D}$. The linkage disequilibrium in a mixed population is effectively the average linkage disequilibrium in the components that are mixed plus the covariance in gene frequencies among the components (Feldman & Christiansen, 1975; Christiansen, 1987, 1999). Applying this result to population $\mathscr{R}$ with immigrants produces the form of the initial linkage disequilibrium difference that appears in the second term of (10).

3. Quality of the approximation

We have evaluated the quality of the approximation (10) by numerical iterations for different sets of selection coefficients in the additive-by-additive epistasis model of (9). Three-, five- and seven-locus models gave similar results, and we report only one pattern of selection for a case of five loci. In this example the coefficients of selection are given by $a_i = -0{\cdot}5\epsilon$, $b_i = -1{\cdot}5\epsilon$ and $c_{ij} = -\epsilon$, and the recipient population is at a monomorphic equilibrium with the 0-allele fixed at all loci (mean fitness equal to $\mu = 1$). The parameter ϵ describing the order of the strength of selection is varied as $\epsilon = 10^{-k}$, $k = 2, 3, 4, 5, 6$ and we will illustrate the strength of selection by the relative difference

$$\omega = \frac{w^+ - w^-}{w^+},$$

where w^- denotes the fitness of the least viable individual in the population and w^+ the highest fitness ($w^+ = 1$ in the example). Note that with $\epsilon = 0{\cdot}01$ selection becomes quite strong against individuals carrying many 1-alleles, in that $w^- = 0$. The relative difference ω varies as $\omega = 10^{-k}$, $k = 0, 1, 2, 3, 4$.

The effect of linkage on the approximation was explored using three different patterns of linkage relationship among the five loci. The first is free recombination in which the loci are unlinked. The remaining two patterns assume the five loci to be placed in order along a chromosome with equal distance between any two neighbouring loci. We assume no interference and consider recombination frequencies of 0·05 and 0·005. Using these parameters we considered three donor populations with the genetic compositions shown in Table 1.

The iterations produced the value of the exact selection barrier B_{exact} given by (4), and we compared it to its approximation B_{approx} given by (10). The comparison is done in terms of a relative error given by

$$\beta = \frac{B_{\text{approx}} - B_{\text{exact}}}{B_{\text{approx}} + B_{\text{exact}}}.$$

If our approximation is accurate we expect a value of β which is on the order of ϵ, and we therefore plot the

https://doi.org/10.1017/S001667230000464X Published online by Cambridge University Press

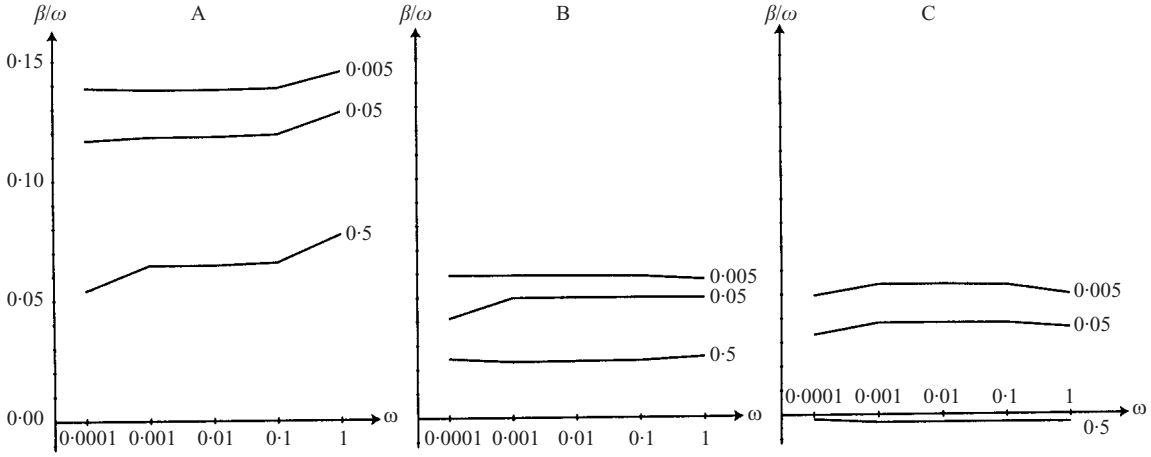

Fig. 1. The normalized relative error of approximation β/ω as a function of the strength of selection ω for five loci. The three drawings correspond to the three donor populations described in Table 1, and the three curves in each drawing correspond to three values of recombination frequencies between neighbouring loci. The additive-by-additive epistasis model is assumed with $a_i = -0{\cdot}5\epsilon$, $b_i = -1{\cdot}5\epsilon$ and $c_{ij} = -\epsilon$ in the recipient population.

normalized value, β/ω, which must be on the order of 1 with respect to ϵ (Fig. 1). The correspondence between the exact barrier and its approximation is very satisfactory because the normalized relative error is virtually independent of ω, and therefore of ϵ. The approximation works even for quite strong selection ($\epsilon = 0{\cdot}01$, $\omega = 1$).

4. Discussion

Zhivotovsky & Christiansen (1995) calculated the strength of the selection barrier (their equation 21) assuming that the differences in gene frequencies between the recipient and donor populations are small, at most on the order of magnitude of the selection coefficients. The barrier strength was on the order of magnitude corresponding to the square of the selection coefficients. Allowing for an arbitrary genetic difference between the two populations, the selection barrier in (10) is on the same order of magnitude as the selection coefficients, and an additional term occurs which can be interpreted in terms of the covariance in gene frequencies for pairs of loci between the recipient and the donor population. For the case of genetically close populations, this term becomes of an order of magnitude lower than the barrier strength and it is therefore neglected. Zhivotovsky & Christiansen's (1995) interpretation of the strength of the selection barrier as having two components still remains, however. For genetically distant populations subject to arbitrary weak viability selection the barrier is given by (8), and the components are given by:

The allelic component: this depends only on the frequencies of alleles in the donor population that are absent in the recipient population.

The linkage disequilibrium component: this depends on both the differences in linkage disequilibria between the recipient population and the donor population, and on the covariance in allele frequencies between these populations.

The allelic component is always positive, but the sign of the linkage disequilibrium component is undetermined. For genetically distant populations the contribution of the product term $\Pi_{i\in K}\Delta p_i$ to the selection barrier can be significant. In particular, if the recipient and donor populations have all alleles in common the allelic component vanishes, and the linkage disequilibrium component becomes the only contributor to the barrier.

The existence of a negative barrier is possible because with multiple loci it is quite unusual that the mean fitness at a stable equilibrium is at a local maximum. With weak selection, the mean fitness increases as the population converges to equilibrium as we may assume that the population rapidly reaches quasi linkage equilibrium (Kimura, 1965; Nagylaki *et al.*, 1999). Introgression is expected to perturb the recipient population away from the linkage equilibrium surface, and the equilibrium mean fitness is maximized only within this surface.

The approximation of the selection barrier in (10) is very satisfactory, and the errors of approximation in the two components of the barrier value are comparable. We can define relative errors β_1 and β_2 respectively for the allelic and the linkage disequilibrium components, and the ratio β_1/β_2 is virtually independent of the strength of selection for the examples given in Fig. 1. For the donor populations A and C the ratio is above but close to one. For donor B the error is larger for the linkage disequilibrium component, giving a ratio of about 0·6.

The mutual barriers between the two populations $\mathscr{R}$ and $\mathscr{D}$ may be studied if we assume both populations to be at a stable genetic equilibrium. With weak

https://doi.org/10.1017/S001667230000464X Published online by Cambridge University Press

selection, the equilibrium values of the linkage disequilibria are of the same order of magnitude as the selection coefficients and therefore negligible compared with the allele-frequency differences. For instance, in the additive-by-additive epistasis model, (10), which describes the barrier that protects population $\mathscr{R}$ against migrants from population $\mathscr{D}$ ($\mathscr{H}_{t+1} = \mathscr{H}_t \times \mathscr{R}$ and $\mathscr{H}_0 = \mathscr{D}$), is

$$B_{\mathscr{D}}^{\mathscr{R}} \approx \frac{1}{\bar{w}^{\mathscr{R}}} \left(\sum_{i \in M} |F_i^{\mathscr{R}}||\Delta p_i^{\mathrm{o}}| - \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1-r_{ij}}{1+r_{ij}} c_{ij}^{\mathscr{R}} \Delta p_i^{\mathrm{o}} \Delta p_j^{\mathrm{o}} \right), \tag{11}$$

and the strength of the selection barrier in population $\mathscr{D}$ against introgression from population $\mathscr{R}$ ($\mathscr{H}_{t+1} = \mathscr{H}_t \times \mathscr{D}$ and $\mathscr{H}_0 = \mathscr{R}$) is

$$B_{\mathscr{R}}^{\mathscr{D}} \approx \frac{1}{\bar{w}^{\mathscr{D}}} \left(\sum_{i \in M} |F_i^{\mathscr{D}}||\Delta p_i^{\mathrm{o}}| - \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1-r_{ij}}{1+r_{ij}} c_{ij}^{\mathscr{D}} \Delta p_i^{\mathrm{o}} \Delta p_j^{\mathrm{o}} \right). \tag{12}$$

We assume that the populations have the same recombination frequencies, and that they are at stable equilibria sufficiently different to make the products of the gene frequency differences $\Delta p_i^{\mathrm{o}} \Delta p_j^{\mathrm{o}}$ of a higher order of magnitude than the selection coefficients. The barrier in a population depends only on the selection in that particular population, and we therefore label the coefficients c_{ij} and F_i by superscripts $\mathscr{R}$ and $\mathscr{D}$. If $c_{ij}^{\mathscr{R}}$ and $c_{ij}^{\mathscr{D}}$ are of the same sign, the linkage disequilibrium component contributes to the two barriers in a similar way, but in general the barrier between the two populations may have different permeability depending on the direction of introgression.

The barrier between genetically close populations at genetic equilibrium where the gene frequency differences are on the same order magnitude as the selection coefficients, is given by

$$B_{\mathscr{D}}^{\mathscr{R}} \approx \frac{1}{\bar{w}^{\mathscr{R}}} \left(\sum_{i \in M} |F_i^{\mathscr{R}}||\Delta p_i^{\mathrm{o}}| - \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1-r_{ij}}{1+r_{ij}} c_{ij}^{\mathscr{R}} \Delta D_{ij}^{\mathrm{o}} \right), \tag{13}$$

and a similar expression for $B_{\mathscr{R}}^{\mathscr{D}}$ (Zhivotovsky & Christiansen, 1995). Thus, (11) provides the relevant terms in the approximation for $|\Delta p_i^{\mathrm{o}}| \gg \epsilon$; the barrier is on the order of magnitude ϵ and the error is of order ϵ^2. Equation (13) gives the approximation for $|\Delta p_i^{\mathrm{o}}| \approx \epsilon$ with the barrier being on the order of magnitude ϵ^2 and the error of order ϵ^3. In the intermediate situation where $|\Delta p_i^{\mathrm{o}}| \gg \epsilon \approx |\Delta p_i^{\mathrm{o}} \Delta p_j^{\mathrm{o}}|$ only the allelic component of the barrier becomes significant unless all loci are polymorphic in both populations, and we then get

$$B_{\mathscr{D}}^{\mathscr{R}} \approx \frac{1}{\bar{w}^{\mathscr{R}}} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1-r_{ij}}{1+r_{ij}} c_{ij}^{\mathscr{R}} (\Delta D_{ij}^{\mathrm{o}} + \Delta p_i^{\mathrm{o}} \Delta p_j^{\mathrm{o}}), \tag{14}$$

which is of the order of magnitude ϵ^2 with an error of order $\epsilon^{5/2}$.

In the calculation of the barrier strength (8), however, we did not assume equilibrium in the donor population $\mathscr{D}$. It only describes the initial immigrants $\mathscr{H}_0$. Thus, the source of the immigrants may as well be a mixture of populations such that the linkage disequilibria in $\mathscr{D} = \mathscr{H}_0$ may be substantial (Feldman & Christiansen, 1975; Christiansen, 1987, 1999). The linkage disequilibria in the donor population may therefore be substantial. The linkage disequilibria at equilibrium in the recipient population may still be neglected, and in the additive-by-additive epistasis model we therefore get the barrier as

$$B_{\mathscr{D}}^{\mathscr{R}} \approx \frac{1}{\bar{w}^{\mathscr{R}}} \left(\left(\sum_{i \in M} |F_i^{\mathscr{R}}||\Delta p_i^{\mathrm{o}}| - \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1-r_{ij}}{1+r_{ij}} c_{ij}^{\mathscr{R}} \right) \times (D_{ij}^{**} + \Delta p_i^{\mathrm{o}} \Delta p_j^{\mathrm{o}}) \right). \tag{15}$$

In case of a mixed $\mathscr{H}_0$, D_{ij}^{**} is related to the gametic linkage disequilibrium between locus i and locus j in the mixture of the donor populations (Christiansen, 1999).

This effect may emerge even if selection works in the same way in all the populations. A multilocus selection model may well have more than one stable equilibrium. For instance, in the two-locus symmetrical viability model of Lewontin & Kojima (1960), a population may for weak selection (loose linkage) possess a stable polymorphic equilibrium and four stable monomorphic equilibria (Christiansen & Feldman, 1975; Feldman & Liberman, 1979; Christiansen, 1999). This allows for a gene frequency difference at both loci between a recipient and a donor population which are both at equilibrium, and this makes both the allelic and the disequilbrium components significant in the barriers in (11) and (12). Even in this simple two-locus model, genetic variation among the components of a mixed donor population is possible, and both of the terms of the linkage disequilibrium components in the barrier of (15) may therefore be significant.

As an illustration we have analysed our results for a simpler mode of selection, namely weak stabilizing selection on a quantitative character. The quantitative trait is determined additively with genotype–environment interaction based on Lerner's model of homeostasis (Zhivotovsky & Feldman, 1992*a*), where a heterozygote at any locus lowers the variance of the environmental contribution to the trait of the individual. Fitness is given in terms of a quadratic deviation from an optimum phenotypic value. The model has the form introduced by Zhivotovsky & Gavrilets (1992) and was analysed by Zhivotovsky & Feldman (1992*a*, *b*). The present analysis follows that of Zhivotovsky & Christiansen (1995). We assume

https://doi.org/10.1017/S001667230000464X Published online by Cambridge University Press

free recombination, that both populations are at a stable equilibrium, and that the recipient population is in a totally polymorphic state. Then the expression for the selection barrier may be simplified to exhibit two terms: one is the squared difference in the mean values of the character, $\Delta x = x^{**} - x^{*}$, and the other is the additive genetic variance, $\Delta V_G = V_G^{**} - V_G^{*}$. Thus,

$$B_{\mathscr{D}}^{\mathscr{R}} \approx \frac{1-\bar{w}^{\mathscr{R}}}{6\bar{w}^{\mathscr{R}} V^{\mathscr{R}}}(\tfrac{1}{2}(\Delta x)^2 + \Delta V_G), \tag{16}$$

where $V^{\mathscr{R}}$ is the phenotypic variance in the recipient population. The first term is always positive and the difference in mean values of the traits therefore always contributes to the barrier. The second term may either increase or decrease barrier strength, in that the population with the greater additive genetic variance will be more susceptible to introgression.

The research was supported in part by grants 94-0163-1 and 9701412 from the Danish Natural Science Research Council, grant 5.18.2023 from the Danish Biotechnological Research and Development Programme and grant 98-04-49292 from the Russian Foundation for Basic Research. We are grateful to Else Løvdal for linguistic revision of the manuscript.

References

Barton, N. H. (1979). The dynamics of hybrid zones. *Heredity* **43**, 341–359.

Barton, N. H. & Bengtsson, B. O. (1986). The barrier to genetic exchange between hybridising populations. *Heredity* **56**, 357–376.

Barton, N. H. & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**, 113–148.

Barton, N. H. & Turelli, M. (1991). Natural and sexual selection on many loci. *Genetics* **127**, 229–255.

Bazykin, A. D. (1969). Hypothetical mechanism of speciation. *Evolution* **23**, 685–687.

Bengtsson, B. O. (1985). The flow of genes through a genetic barrier. In *Evolution: Essays in Honour of John Maynard Smith* (ed. J. J. Greenwood & M. Slatkin), pp. 31–42. Cambridge: Cambridge University Press.

Christiansen, F. B. (1987). The deviation from linkage equilibrium with multiple loci varying in a stepping-stone cline. *Journal of Genetics* **66**, 45–67.

Christiansen, F. B. (1999). *Population Genetics of Multiple Loci*. Chichester: Wiley.

Christiansen, F. B. & Feldman, M. W. (1975). Selection in complex genetic systems. IV. Multiple alleles and interactions between loci. *Journal of Mathematical Biology* **2**, 179–204.

Christiansen, F. B., Andreasen, V. & Poulsen, E. T. (1995). Genotypic proportions in hybrid zones. *Journal of Mathematical Biology* **33**, 225–249.

Feldman, M. W. & Christiansen, F. B. (1975). The effect of population subdivision on two loci without selection. *Genetical Research* **24**, 151–162.

Feldman, M. W. & Liberman, U. (1979). On the number of stable equilibria and the simultaneous stability of fixation and polymorphism in two locus models. *Genetics* **92**, 1355–1360.

Kimura, M. (1965). Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. *Genetics* **52**, 875–890.

Lewontin, R. C. & Kojima, K. (1960). The evolutionary dynamics of complex polymorphism. *Evolution* **14**, 458–472.

Nagylaki, T. (1992). *Introduction to Theoretical Population Genetics*. *Biomathematics*, vol. 21. Berlin: Springer.

Nagylaki, T., Hofbauer, J. & Brunovski, P. (1999). Convergence of multilocus systems under weak epistasis or weak selection. *Journal of Mathematical Biology* **38**, 103–133.

Slatkin, M. (1972). On treating the chromosome as the unit of selection. *Genetics* **72**, 157–168.

Turelli, M. & Barton, N. H. (1994). Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? *Genetics* **138**, 913–941.

Zhivotovsky, L. A. & Christiansen, F. B. (1995). The selection barrier between populations subject to stabilizing selection. *Evolution* **49**, 490–501.

Zhivotovsky, L. A. & Feldman, M. W. (1992*a*). On models of quantitative genetic variability: a stabilizing selection–balance model. *Genetics* **130**, 947–955.

Zhivotovsky, L. A. & Feldman, M. W. (1992*b*). On the difference between the mean and optimum of quantitative characters under selection. *Evolution* **46**, 1574–1578.

Zhivotovsky, L. A. & Gavrilets, S. (1992). Quantitative variability and multilocus polymorphism under epistatic selection. *Theoretical Population Biology* **42**, 254–283.

Zhivotovsky, L. A. & Pylkov, K. V. (1998). A general multilocus two-allele model for epistatic selection. *Journal of Genetics* **77**, 115–121.