Polymorphism in heterogeneous environments, evolution of habitat selection and sympatric speciation: soft and hard selection models

T. DE MEEÛS*, Y. MICHALAKIS‡, F. RENAUD and I. OLIVIERI‡

Laboratoire de Parasitologie Comparée, URA 698 CNRS, Université Montpellier II (USTL), Place E. Bataillon, 34095 Montpellier Cedex 05, France

Summary

The adaptation to a variable environment has been studied within soft and hard selection frameworks. It is shown that an epistatically determined habitat preference, following a Markovian process, always leads to the maintenance of an adaptive polymorphism, in a soft selection context. Although local mating does not alter the conditions for polymorphism maintenance, it is shown that, in that case, habitat selection also leads to the evolution of isolated reproductive units within each available habitat. Habitat selection, however, cannot evolve in the total absence of adaptive polymorphism. This represents a theoretical problem for all models assuming habitat selection to be an initially fixed trait, and means that within a soft selection framework, all the available habitats will be exploited, even the less favotirable ones.

On the other hand, polymorphism cannot be maintained when selection is hard, even when all individuals select their habitat. Here, the evolution of habitat selection does not need any prerequisite polymorphism, and always leads to the exploitation of only one habitat by the most specialized genotype. It appears then that hard selection can account for the existence of empty habitat and for an easier evolution of habitat specialization.

Keywords: polymorphism; variable environments; habitat selection; speciation; soft selection; hard selection

Introduction

Over the last forty years, evolutionary biologists have debated three important questions concerning the maintenance of genetic polymorphism: (1) how can selection in a variable environment maintain polymorphism; (2) what role can habitat selection play and (3) does this render possible the evolution of reproductive isolation between two sympatric entities? A key factor in this debate is the distinction between soft and hard selection (Wallace, 1968). Wallace (1975) defined selection as soft when it is both frequency and density dependent and as hard when it is both frequency and density independent.

As far as genetic polymorphism in a variable environment is concerned, the soft and hard selection terms have respectively been applied to Levene's (1953) and Dempster's (1955) models. In Levene's model, population regulation occurs independently within each habitat (constant number of individuals produced by each habitat, at each generation). This allows density and frequency dependent selection to take place within the habitats (Barton and Clark, 1990). In

---

* To whom correspondence should be addressed.

‡ Current address: Station INRA d'Amélioration des Plantes de Montpellier, Domaine de Melgueil, 34130 Mauguio, France.

Dempster's model, population regulation occurs before habitat colonization (constant number of colonizators). This leads to a density and frequency independent kind of selection (despite the fact that regulation indeed occurs).

In Dempster's model (hard selection) the conditions for the maintenance of polymorphism are the same as those in a homogeneous environment (i.e. overdominance relative to the arithmetic mean). In Levene's (soft selection) model these conditions are much wider (i.e. overdominance relative to the harmonic mean). In particular, polymorphism cannot be maintained in Dempster's model when one allele is dominant, while it may in Levene's model (Maynard-Smith, 1962).

Nevertheless, the conditions for maintenance of polymorphism in soft selection models correspond to very severe conditions (Maynard-Smith, 1962, 1966; Maynard-Smith and Hoekstra, 1980).

In such models, environmentally induced habitat selection (e.g., females that tend to lay more eggs in habitats where they were born) broadens the conditions of maintenance of polymorphism (Maynard-Smith, 1966, t970). But, here again, the selective advantages must remain high (Hoekstra *et al.,* 1985).

Templeton and Rothman (1981), Rausher (1984), Garcia-Dorado (1986, 1987) and Hedrick (1990a) have demonstrated that the conditions required for the maintenance of polymorphism are considerably broadened if habitat selection is genetically determined by adaptive gene pleiotropy (the same genes determine adaptation and habitat selection). In certain cases, no selective advantage is necessary. No clear assumption is made, by these authors, concerning the origin and evolution of such habitat selection.

These problems are directly linked to the theory of sympatric speciation. This is a subject of some controversy, involving numerous authors (e.g., Mayr, 1963, 1982; Maynard-Smith, 1966; Bush, 1975; Futuyma and Mayer, 1980; Felsenstein, 1981; Tauber and Tauber, 1989; Diehl and Bush, 1989). The most convincing population genetics models leading to the sympatric evolution of two species have to involve a previous stable polymorphism and/or the existence of some kind of habitat selection (e.g., Maynard-Smith, 1966; Rice, 1984, 1987). Here again, though, the evolution of habitat selection *per se* is not described.

In this paper, a new theoretical approach to the simultaneous study of the evolution of adaptation and habitat selection is proposed. A two-allele, two-independent-loci model is proposed in both soft and hard selection frameworks. The first locus determines a better adaptation to one, of the two, available habitats. The second locus controls the preference or the indifference of individuals to the most favourable habitat. There is thus an epistatic influence of the adaptive locus upon the habitat choice allele. Habitat selection, inspired from Doyle (1975), follows a Markovian process where the two states are the two habitats, the most favourable one being absorbent. The dominant case is assumed for both loci for the sake of simplicity and to exclude overdominance from the conclusions. Overdominance is indeed irrelevant to the study of specialization and speciation mechanisms. Also, the local mating hypothesis is adopted (panmixia within each habitat, independently). This does not change the conditions for polymorphism in the case of dominance of one adaptive allele (Strobeck, 1974), and renders possible the study of the possible evolution of two genetic entities (speciation).

In this paper, the different possible equilibria for each of the two loci, for each kind of population regulation (soft or hard), have been studied. The influence of the different parameters of the model (adaptation, cost of dispersal, intensity of choice, etc.) on the genotypic structure of the adaptation locus have been analysed, both in the case of habitat selection and in the case of indifference. The information gained from this approach has been used to discuss the prerequisite conditions for the emergence of different species. Finally, the dramatic differences found between hard and soft selection models are discussed. These differences could indeed help

the interpretation of patterns found in natural populations, as far as the regulation mode is involved.

The general framework of the models

A species exists as a metapopulation in a landscape with two possible habitats 1 and 2 with frequencies c and $1 - c$, respectively. These habitats could be different plant species for herbivores or different host species for parasites. For simplicity, let the generations be discrete and not overlapping. Adults reproduce at random within each subpopulation (local mating). Panmixia occurs only locally.

THE ADAPTATION IN THE TWO AVAILABLE HABITATS

A first locus, with two alleles A and a, determines the viability in each habitat in the following manner (Table 1). A is assumed dominant, and individuals carrying it survive better in habitat 1. Individuals of genotype aa do better in habitat 2 (*sl £ x, s2 £y*, Table 1). Only those cases where *sl £ y* (aa less fit than AA and Aa in habitat 1) and $s2 \leq x$ (AA and Aa less fit than aa in habitat 2) are considered.

Table 1. Fitness of the different genotypes

	In habitat 1	In habitat 2
AA or Aa	1 - sl	1 - x
aa	1 - y	1 - *s2*

THE CHOICE OF THE MOST FAVOURABLE HABITAT

Let us consider a second locus with two alleles, B and b, 'B being dominant This locus is independent of the previous one (recombination rate = 1/2). Individuals of genotype bb do not choose at all, and settle in the first habitat they encounter. Individuals of genotypes B refuse to settle until they encounter the habitat in which they best survive according to their genotype at the first locus. They cannot, however, refuse an unfavourable habitat for more than a given number of tries, which is i for genotypes AB and j for genotypes aaB: at the ith (or j^{th}) try, individuals of genotypes B do not choose any more and settle in the habitat where they happen to be. Whenever an individual moves from one patch to another, it has a survival probability d. Habitat selection can thus be costly ($0 < d \leq 1$).

The process of finding and settling is then a Markovian process (as in Doyle's model, 1975) where the two states are the two habitats, the most favourable one being absorbent and the other one being repulsive.

If P_{AB1} and P_{AB2} are the probability of settling in habitat 1 and 2 respectively, for AB, we then have:

$$P_{AB1} = cd + cd(1 - c)d + cd[(1 - c)d]^2 + \ldots + cd[(1 - c)d]^{(i-1)}$$

or

$$P_{AB1} = (cd)\{1 - [(1 - c)d]^i\}\{1 - (1 - c)d\}^{-1}$$

and

$P_{AB2} = [(1 - c)d]^i$

Similarly, for the genotypes aaB, one obtains:

$P_{aB1} = (cd)^j$,

$P_{aB2} = (1 - c)d[1 - (cd)^j][1 - (cd)]^{-1}$

Note that, for high values of i and j, $P_{AB1} = cd[1 - (1 - c)d]^{-1}$, $P_{AB2} \approx 0$, $P_{aB1} \approx 0$ and $P_{aB2} \approx (1 - c)d[1 - cd]^{-1}$. This is why i and j will be considered as big enough in all the following, unless when specified otherwise.

GENOTYPIC FREQUENCIES IN EACH HABITAT, AFTER COLONIZATION AND SELECTION

Genotypes which do not choose (bb)

Let $G(x)$ be the initial frequency of genotype x in the propagule pool.

Among the d survivors, a proportion c reaches habitat 1 and $1 - c$ habitat 2. In these habitats, selection occurs in the way described in Table 1. Then the genotypic frequencies ($Gi(x)$) in each habitat are thus as follows

in habitat 1;

$G1(AAbb) = cd(1 - s1)G(AAbb)(W1)^{-1}$

$G1(Aabb) = cd(1 - s1)G(Aabb)(W1)^{-1}$

$G1(aabb) = cd(1 - y)G(aabb)(W1)^{-1}$

in habitat 2;

$G2(AAbb) = (1-c)d(1 - x)G(AAbb)(W2)^{-1}$

$G2(Aabb) = (1-c)d(1 - x)G(Aabb)(W2)^{-1}$

$G2(aabb) = (1-c)d(1 - s2)G(aabb)(W2)^{-1}$

where $W1$ and $W2$ represent the relative numbers of colonizators inhabiting each habitat after dispersal, settlement and selection.

Genotypes which choose (BB and Bb)

in habitat 1;

$G1(AABB) = P_{AB1}(1 - s1)G(AABB)(W1)^{-1}$

$G1(AABb) = P_{AB1}(1 - s1)G(AABb)(W1)^{-1}$

$G1(AaBB) = P_{AB1}(1 - s1)G(AaBB)(W1)^{-1}$

$G1(AaBb) = P_{AB1}(1 - s1)G(AaBb)(W1)^{-1}$

$G1(aaBB) = P_{aB1}(1 - y)G(aaBB)(W1)^{-1}$

$G1(aaBb) = P_{aB1}(1 - y)G(aaBb)(W1)^{-1}$

in habitat 2;

$G2(AABB) = P_{AB2}(1 - x)G(AABB)(W2)^{-1}$

$G2(AABb) = P_{AB2}(1 - x)G(AABb)(W2)^{-1}$

$G2(AaBB) = P_{AB2}(1 - x)G(AaBB)(W2)^{-1}$

$G2(AaBb) = P_{AB2}(1 - x)G(AaBb)(W2)^{-1}$

$G2(aaBB) = P_{aB2}(1 - s2)G(aaBB)(W2)^{-1}$

$G2(aaBb) = P_{aB2}(1 - s2)G(aaBb)(W2)^{-1}$

If $w_i(x) = $ (survival in habitat i of genotype x) (probability for x to colonize i), then:

$W_i = \Sigma_x[G(x)w_i(x)]$

These surviving adults then mate at random, independently within each habitat. The genotypic

frequencies among offspring, $G'1(x)$ and $G'2(x)$, can hence be calculated for each habitat, giving a recombination rate of 1/2 (independence).

NEXT GENERATION PROPAGULE POOL

In the soft selection model

Here we consider the propagule production in each site to be constant and shared by the individuals exploiting it (density-dependent propagule production). Let us assume that the relative production of each habitat is c for habitat I and I - c for habitat 2 (i.e., equal to habitat frequencies or sizes). Thus, the overall frequency $G'(x)$ of a given genotype x in the propagule pool of the next generation, is the arithmetic mean of the within habitat genotypic frequencies: $G'(x) = cG'1(x) + (1 - c)G'2(x)$.

In the hard selection model

In our approach, the differences between a hard and a soft selection model concern only the propagule production of each habitat. The number of propagules is equal to WI for habitat 1, and to W2 for habitat 2. Following this hypothesis, the frequency of genotype x in the propagule pool of the next generation becomes:

$$G'(x) = \{G'1(x)W1 + G'2(x)W2\}\{W1 + W2\}^{-1}$$

The propagules then recolonize the two habitats to produce a new generation following the rules described in the previous sections.

Although not very high, the complexity of this model is such that any general solution appears analytically intractable. Particular cases can, however, be solved analytically using the protected polymorphism technique (both alleles must increase in frequency when rare), which gives necessary and sufficient conditions (Gliddon and Strobeck, 1975). Otherwise, a program written in Turbo-Pascal (Borland), enabling numerical simulations, is used to explore this model and find some particular solutions for several kinds of parameter sets. The parameters are the following: d, the proportion of surviving propagules when moving from one habitat to another; x, $s1$, y and $s2$, the adaptive parameters (see Table 1); c, the habitat 1 frequency; i and j, the maximum numbers of tries for phenotypes [AB] and [aaB] respectively; and $G0(x)$, genotype x frequency in the initial propagule pool.

Results

EVOLUTION OF LOCUS A/a (ADAPTATION LOCUS) WITH NO HABITAT SELECTION

In the soft selection model

This corresponds to Levene's (1953) multiple niche model or, more exactly, to Maynard-Smith's (1962) and Prout's (1968) models, because there are only two niches and A is dominant. Within our framework, these conditions write as follows:

A is protected if $c > (x - s2)(1 - y)[(1 - s1)(1 - s2) - (1 - x)(1 - y)]^{-1}$ and
a is protected if $c < (x - s2)(1 - s1)[(1 - s1)(1 - s2) - (1 - x)(1 - y)]^{-1}$

In the hard selection model. This corresponds to Dempster's (1955) model, with local mating instead of panmixia. Polymorphism cannot be maintained in the case of a dominant allele (Maynard-Smith and Hoekstra, 1980). Let p and p' be the frequencies of A at generation t and $t + 1$ respectively. Here, no stable polymorphism is possible since $p' - p$ has the same sign as:

$$c[(1 - s1) - (1 - y)] - (1 - c)[(1 - s2) - (1 - x)]$$

which does not depend on p. This provides a generalization of the fact that local mating does not change the conditions for protected polymorphism (demonstrated by Strobeck in 1974 for a soft selection model).

EVOLUTION OF Locus A/a (ADAPTATION LOCUS), WHEN THERE IS ONLY HABITAT SELECTION

In the soft selection model
In this situation, a stable polymorphism is always reached at the adaptive locus, as long as $i > 1$ and $j > 1$ (Appendix A). Here, the rarest, which only settle in the emptiest habitat is indeed highly favoured. With local mating, there is no beterozygote Aa at equilibrium, providing i and j are large enough. The equilibrium frequency of A is thus equal to the frequency of habitat 1, i.e., c.

In the hard selection model. Appendix A shows that the sign of $p' - p$ (the evolution of the frequency of allele A, from one generation to the next) depends only on the expression:

$$(1 - s1)c[1 - (1 - c)d]^{-1} - (1 - s2)(1 - c)[1 - cd]^{-1}$$

So, when B (choosy allele) is fixed, the evolution at the adaptation locus is independent of the allelic frequencies. This means that no stable polymorphism is possible.
Let $W_{AB} = (1 - s1)c[1 - (1 - c)d]^{-1}$ and $W_{aB} = (1 - s2)(1 - c)[1 - cd]^{-1}$. Then, if $W_{AB} > W_{aB}$, A becomes fixed, if $W_{AB} < W_{aB}$, a becomes fixed and if $W_{AB} - W_{aB}$, then the situation is neutral.
Particularly, if dispersal survival is high ($d \gg 1$) then $W_{AB} \approx (1 - s1)$ and $W_{aB} \approx (1 - s2)$ and all depends on the adaptive parameters; if the dispersal survival is low ($d \approx 0$), then $W_{AB} \approx c(1 - s1)$ and $W_{aB} \approx (1 - c)(1 - s2)$ and habitat frequencies and adaptive parameters are equally important. If $s1 \approx s2$ (AA and Aa are as fit in habitat 1 than aa in habitat 2) the sign of $p' - p$ depends only on the sign of $c - 1/2$; here the genotypes that choose the largest habitat will spread (except when $d=1$ in which case $p' - p = 0$).

EVOLUTION OF THE LOCUS B/b (HABITAT SELECTION) WHEN THE A/a LOCUS (ADAPTATION) IS MONOMORPHIC: (PHENOTYPES [B] CHOOSE ONLY ONE HABITAT)

In the soft selection model

In these conditions, choice is never selected for, and allele b spreads (Appendix B). Here, because of the density dependent regulation, refusing to settle in an empty habitat is always very unfavourable.

In the hard selection model

Let A be the fixed allele at the adaptive locus. Let q and q' be the allelic frequencies of B in generations t and $t + 1$, respectively. In Appendix B it can be seen that the sign of $q' - q$ depends exclusively on the sign of:

$$(1 - s1)cd[1 - (1 - c)d]^{-1} - (1 - x)$$

which is frequency independent (no stable polymorphism can be found) and the conditio for q to increase (B becomes fixed) is:

$$(1 - s1)cd[1 - (1 - c)d] > (1 - x)$$

otherwise B is eliminated or neutral.

Symmetrically, the condition for B to be selected when allele a is initially fixed is:

$$(1 - s2)(1 - c)d[1 - cd]^{-1} > (1 - y)$$

If d is large enough (d $\approx$ 1) the evolution of B becomes independent of c (habitat frequency). On the other hand, when d becomes low, the frequency of the best habitat and the survival in it have to increase considerably compared with those of the other habitat. If habitat 1 and habitat 2 are identical for A ($s1 - x$) or for a ($s2 = y$), i.e., when the environment is homogeneous, then B is always eliminated, except when d=1 in which case locus B/b is neutral. The ability of the choosy allele to be fixed here is a very important difference between this model and the soft selection one.

SIMULTANEOUS EVOLUTTON OF Loci B/b (HABITAT SELECTION) AND A/a (ADAPFATION)

Here, most of the results come from computer simulations using our Turbo-Pascal program. Some particular cases, however, are solved analytically. Simulations show that a polymorphism cannot be selected for at the choice locus in the soft selection model and at either of the two loci in the hard selection one. Moreover, the two loci are independent and local matings occur at random. Thus, linkage disequilibria always quickly reach low values. This is why, for the sake of simplicity, linkage disequilibria are neglected when analytical solutions are explored.

In the soft selection model

Protection of B (a) 1.1. Choice is costly ($d < 1$): From section 3, it is known that the existence of an adaptive polymorphism is critical for B to be selected for in a soft selection framework. On the other hand, when A/a is polymorphic, choice can be selected for. Figure 1 shows conditions where B is protected, for different values of the adaptive parameters and different values of d. These are obtained assuming (1) high values for i and j (number of tries); (2) $c = 1 - c = 1/2$ (habitat frequencies); (3) $s1 = s2 = 0$ and $x = y$ (adaptive parameters, see Table 1); (4) $p0 = 1/2$ (starting frequency of A); and (5) global random mating and linkage equilibrium
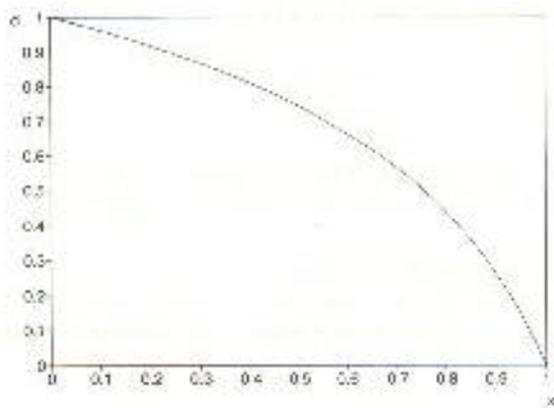

Figure 1. Numerical application illustrating the influence of x and y (selective costs in the wrong habitat) and of d (dispersal survival rate) on the evolution of B/b (habitat choice locus). Here x = y, $c = 1/2$ (habitat frequency), $s1 = s2 = 0$ (costs in the best habitat), $p0 = 1/2$ (initial frequency of A), and i and j (number of allowed tries for choosy genotypes) are infinite. Panmixia and linkage equilibrium are assumed for simplicity. Selection is soft. The curve corresponds to $d = \{2(1 - x) [5 + 3(1 - x)]\} \{3 + (1 - x)[10 + 3(1 - x)]\}^{-1}$. Above this curve, B spreads. Beneath the curve, B is eliminated.

(enabling analytical treatment). In such conditions, adaptive polymorphism is always protected and B is protected if

$$d > \{2(1 - x)[5 + 3(1 - x)]\} \{3 + (1 - x) [10 + 3(1 - x)]\}^{-1}$$

It should be noted that, despite such a favourable set of parameters, less conditions exist that lead to the protection of B (above the curve of Fig. 1) compared with those that lead to its elimination. If the parameter set does not allow for a protected adaptive polymorphism, computer simulations show that the selection for B becomes very uneasy. It was shown in section 2, however, that, if B is protected, so are A and a.

(a) 1.2. Choice is not costly ($d=1$): This special case is closer to classical population genetics models dealing with polymorphism with habitat selection (e.g. Maynard-Smith, 1962, 1966; Rausher, 1984; Garcia-Dorado, 1987 or Hedrick, 1990a).

In order to simplify the algebra, it is assumed that there is a linkage equilibrium between the two loci, within a haploid population. If choice is not costly, and if i and j are large, it can be shown that the frequency of B, after settlement and selection, is equal to

$$q1 = pq(1 - s1)(W1)^{-1} \text{ in habitat 1, and}$$
$$q2 = (1 - p)q(1 - s2)(W2)^{-1} \text{ in habitat 2}$$

where p and q are the respective frequencies of A and B at generation t and where

$$W1 = p(1 - s1)[q + (1 - q)c] + (1 - p)(1 - q)c(1 - y) \text{ and}$$
$$W2 = (1 - p)(1 - s2)[q + (1 - q)(1 - c)] + p(1 - q)(1 - c)(1 - x)$$

Within a soft selection framework, the frequency of B in the next generation is equal to $q' = cq1 + (1 - c)q2$. The evolution of B is given by the sign of $q' - q$. Substituting $q1$ and $q2$ by their respective values, it can be shown that $q' - q$ is positive if

$$W_{R1}(a)/W_{R2}(a) + q[RW_{R1}(a) + 1/(RW_{R2}(a))] < 1$$

where $W_{R1}(a) = (1 - y)/(1 - s1)$ and $W_{R2}(a) = (1 - s2)/(1 - x)$ are the relative fitness of a, compared with A, in habitats 1 and 2, respectively, and $R = c(1 - p)/[(1 - c)p]$ is the ratio, between a and A, of settlements in the wrong habitat while not choosing.

It is easily seen that if p (frequency of A) is too close to 0 or I this inequality cannot be true. When p lies between these two limits, however, B may spread, providing that polymorphism stays long enough. The more the conditions will be close to those of protected polymorphism, the longer B will be selected for. As suggested by Maynard-Smith (1962), a protected polymorphism indeed represents a sufficient condition for B to invade the population, but, as shown here, not a necessary one.

It can be concluded that, in a soft selection model, choosing is favourable in less restrictive conditions than those for protected polymorphism, as long as choosing is not costly. These conditions, however, become very restrictive as soon as habitat choice becomes more realistically costly.

Some special cases (1) Influence of the number of tries (i and j): varying i and j, independently, allows us to quantify the respective influence of A or a in the selection for B. The results obtained in the following numerical simulations illustrate this point (Fig. 2). If we assume that:

$$d = 0.8 \text{ and } c = (1 - c) = 1/2,$$
$$s1 = 0, x = 0.8 \text{ and } s2 = y = 0.4,$$
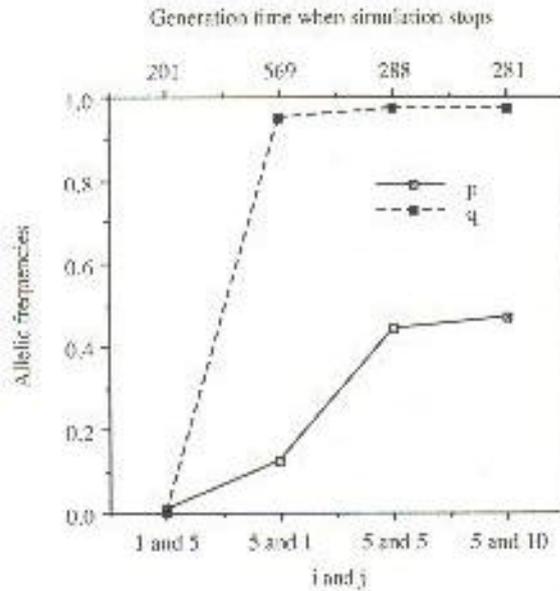$$p0 = 0.5 \text{ and } q0 = 0.25 \text{ with initial linkage equilibrium.}$$

Figure 2. Results of simulations obtained varying *i* and *j* independently. Here, AA and Aa are highly specialized for habitat 1 *(sl = 0, x = 0.8)*, while aa survive as well in both habitat *(s2 = y = 0.4)*. Dispersal survival is high *(d = 0.8)*, habitats have the same size *(c = 0.5)* and simulations stop when the difference between two generations is less than 0.0001 for the frequencies of A *(p)* and B *(q)*. Selection is soft. Note that aa select for B faster when faced with a more choosy competitor (increasing *i*), but that this leads a to lower equilibrium frequencies.

It is then easy to see that genotypes AA and Aa are highly specialized in habitat 1 and that genotypes aa survive as well as habitats 1 and 2. Varying *i* (number of allowed tries for AB genotypes) and *j* (number of allowed tries for aaB genotypes) gives the following results (the end of simulations takes place when $p' - p \leq 0.0001$ or $q' - q \leq 0.0001$). In Fig. 2, it can be seen that when genotypes AB are more choosy (i.e. larger i) aa consequently select B for more strongly and/or faster.

When a specialized phenotype competes with a less specialized one, the specialist will occupy most of the habitat it chooses. The generalist genotypes do not survive much better in one or the other habitat. But, to reproduce in the emptiest habitat, is more favourable than to do so in the specialist's one, which is crowded. This leads the generalist to select B for. In other words, the competition with a specialist obliges the generalist to become a specialist of the less occupied habitat. This recalls the results of Rosenzweig (1978, 1987) and those of Wilson and Turelli (1986). Thus, the influence of each adaptive allele can be tested by varying *i* and *j*. Notice that despite the fact that genotypes aa select B for (by their own), this leads them to a lower equilibrium frequency. This comes from the dominance of allele A.

(2) Consequence of the dominance of A: in the simulations presented in Fig. 2, *i* and *j* are not infinite. Thus, some heterozygotes are left at the end.

In the case of stable equilibria on A/a, since Aa survive better in habitat 1 than in habitat 2, allele a will be better represented in habitat 1 than A in habitat 2. Hence allele a will have a higher equilibrium frequency. Indeed, if we take, for example, *x = y, sl = s2* and *c = 1/2*, at the equilibrium, the frequencies of phenotypes [A] and [aa] will be equal. Thus, the frequency

of the genotype aa will be equal to that of AA + Aa. It is clear that the equilibrium frequency of a will be greater than that of A.

This phenomenon has an interesting consequence. Allele a will have a lower equilibrium frequency when aa is more choosy (large j) than when aa is less choosy (low j) as shown in the simulations presented in this section (Fig. 2). This is because choice lowers the heterozygote frequency.

(3) Evolution of locus B/b (habitat selection) in a constant environment ($x = y = s1 = s2$): in this situation, the genotypes B choose habitat 1 or 2, depending on their genotype at locus A/ a, but, apart from density-dependent effects occurring in the soft selection model, their viability once established does not depend on where they happen to settle.

As shown below, whether a polymorphism occurs or whichever allele invades, is strongly dependent upon linkage disequilibrium between the two loci. Two cases have to be distinguished. (1) Habitat selection is costly ($d < 1$). If the dispersal survival is less than unity, allele b (no habitat selection) always becomes fixed, whatever $i > 1$ and $j > 1$, as shown by all the computer simulations we have carried through. (2) No cost to habitat selection ($d = 1$). Let us consider that propagules [AB] settle only in habitat 1 and propagules [aaB] only in habitat 2 (large i and j). Let q be the frequency of allele B in the pool of propagules in a given generation, and $q1$ and $q2$ its frequencies in habitat 1 and 2, respectively. Then, if q' is its frequency in the next generation, we have (Appendix C):

$$q' - q = (W1 - c)(q2 - q1),$$

where $W1$ represents the relative number of viable colonizators in habitat 1:

$$WI = \Sigma_x G(x) w_i(x).$$

The equilibrium is obtained for $W1 = c$ or $q1 = q2$. At equilibrium, linkage disequilibrium can be neglected. Then, the condition for an (unstable) equilibrium for locus B/b is: $G(aa) = 1 - c$ (Appendix C). This equilibrium frequency is such that the probability of settling into habitat 1 or 2 is equal to c and $1 - c$, respectively, for each of the three genotypes BB, Bb and bb. This explains why B is then neutral. If, however, $G(aa)$ is moved away from ($1 - c$), then B will be more likely selected against, as only genotypes bb keep these optimal probabilities of landing in habitat I or 2. When G(aa) reaches its equilibrium value again, B is once more neutral. It can happen, however, that B is eliminated before A reaches its equilibrium frequency, in which case A becomes neutral. A numerical example is given in Fig. 3. The frequency of B influences the distribution of A between the two habitats: the higher q, the bigger the difference between $p1$ (frequency of A in habitat 1) and $p2$ (the frequency of A in habitat 2).

Numerical applications show that, in certain conditions of genotypic structure, B may increase in frequency. This is because the dominance of B allows a given habitat to become crowded without being crowded with B (i.e., when there is a majority of heterozygotes). Let us imagine, for example, a population initially composed with AABB and aaBB individuals only, with $G(AABb) >> G(aaBB)$. Then, after settlement we will have $q2 >> q1$, and $W1 = G(AABB)$ close to 1. $c \neq 1$, thus $W1 > c$. So $q' - q = (W1 - c)(q2 - q1) > 0$, and B will increase in frequency till the next equilibrium is reached ($G(aa) = 1 - c$).

It is shown in Appendix C that, in a haploid model, these conditions could never be realized (B selected against or neutral). However, an unstable equilibrium is reached when the frequency of A is $p = c$, that is an equal probability for B and b to settle in habitat 1 or 2. This represents a similar conclusion to the one obtained with the diploid assumption. This similarity is in agreement with Gliddon and Strobeck's (1975) results.
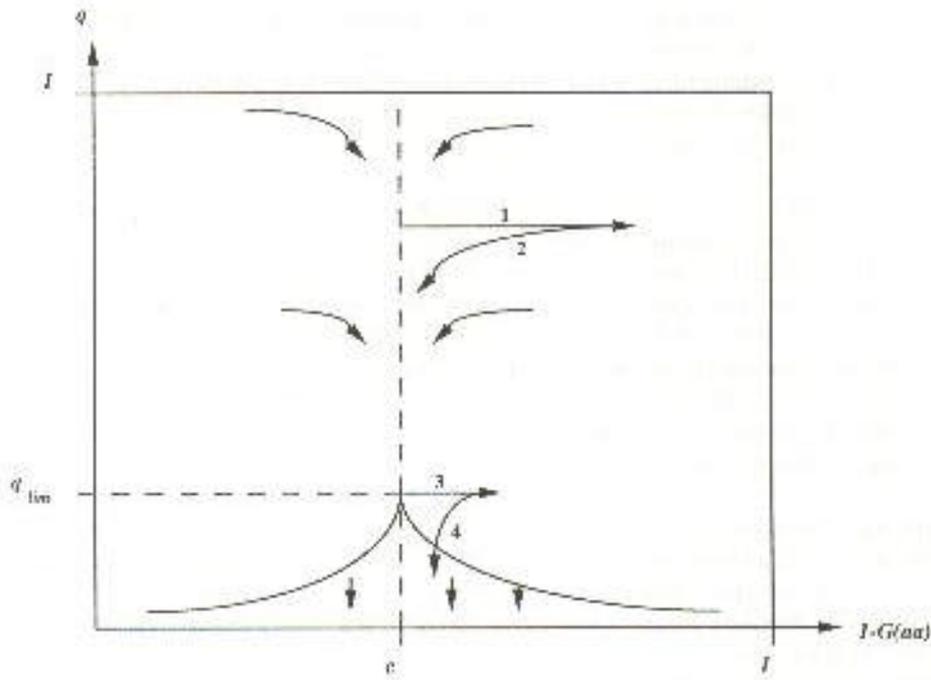
Figure 3. Evolution of habitat selection (locus B/b) and of adaptation (A/a) when selection is soft, the environment is homogeneous (neutrality) and when choosing is not costly ($d = 1$). Under a frequency q_{lim} of B, A/a is neutral and B is selected against. Above this limit ($q > q_{lim}$) A reaches its equilibrium frequency $G(aa) = (1 - c)$, and then B is neutral. A disturbance of type 1 will lead to a slight decrease of q (trajectory 2), until $G(aa) = 1 - c$ again. A disturbance of type 3 leads to trajectory 4, leading to the elimination of B, in which case A/a becomes neutral.

In the hard selection model

Evolution at both loci Different cases can be considered using the results obtained in the previous sections for this model. Let us consider that $P_{AB1} = cd[1 - (1 - c)d]^{-1}$ and $P_{aB2} = (1 - c)d[1 - cd]^{-1}$. P_{AB1} and P_{aB2} are the settlement rates of propagules AB and aaB, respectively (i.e., the proportion of choosy genotypes able to settle in the habitat they prefer). If

> (1) $(1 - s1)P_{AB1} > (1 - x)$, then B is favoured with A.
> (2) $(1 - s2)P_{aB2} > (1 - y)$, then B is favoured with a.
> (3) $(1 - s1)P_{AB1} > (1 - s2)P_{aB2}$, then A is favoured with B.
> (4) $(1 - s1)P_{AB1} < (1 - s2)P_{aB2}$, then a is favoured with B.
> (5) $c(1 - s1) + (1 - c)(1 - x) > (1 - c)(1 - s2) + c(1 - y)$ then A is favoured with b.
> (6) $c(1 - s1) + (1 - c)(1 - x) < (1 - c)(1 - s2) + c(1 - y)$, then a is favoured with b.

Then

> (i) If (1), (2) and (3) are true, choosing and exploiting habitat 1 is favourable and A and B are selected for together.
> (ii) If (1), (2) and (4) are true, choosing and exploiting habitat 2 is favourable and a and B are selected for together.

(iii) If (1) and (3) are true but (2) is false, choosing is favourable only in the presence of a sufficient quantity of allele A.

   If A is initially sufficiently present, B is selected for with A.

   If A is not sufficiently present, or if B is associated in majority with a, B is selected against. In this case, B may be eliminated, depending on the initial frequencies, and A or a is selected for if (5) or (6) is true, respectively.

(iv) If (2) and (4) are true but (3) is false, choosing is favourable only in the presence of a sufficient quantity of allele a.

   If a is initially sufficiently present B is selected for with a.

   If a is not sufficiently present, or if B is associated in majority with A, B is selected against. In this case B may be eliminated, depending on the initial frequencies, and A, or a, spreads if (5) or (6) is true, respectively.

(v) If (1) and (2) are false then choosing is unfavourable and B is eliminated with the adaptive aliele (A or a), exploiting the less favourable habitat (depending on *c* and within habitat survival parameters).

If AA and Aa exploit habitat I as efficiently as aa exploits habitat 2 ($sl = s2$) and $x = y$) and if habitat frequencies are identical (c = 1/2), then the adaptive allele, which is initially the most associated with B, is selected against by hitch-hiking. This may lead to the elimination of this allele if B is not eliminated beforehand (depending on the initial frequencies). Roughly, B becomes fixed with the most efficient adaptive allele which exploits the most frequent habitat; but the latter (adaptation allele) needs the former (B) to be selected for and vice versa.

It is worth noting that polymorphism at the habitat choice locus can be maintained neither in the soft selection nor in the hard selection model. All the simulations we have carried through always end with the fixation of B or b (except when neutral).

The role played by i and j (maximum number of allowed tries)   In all the previous parts we have neglected the role of these parameters in the hard selection model, assuming them to have sufficiently high values. For low values of i and j, however, things may change. Here the algebra becomes much more complex. Nevertheless, simulations show that the most choosy genotype (*i* < *j* or j < i) will get a selective advantage when choice is favourable, and be disadvantaged if not.

From sections 2, 3 and 4, we know that, when the survival is high ($d \approx 1$), the evolution of alleles A and a, when B is fixed (section 2), or the evolution of B and b, when A (or a) is fixed (sections 3 and 4), become independent of habitat frequency. These points do not hold when *i* or *j* have low values (falling in the most favourable habitat becomes more dependent on the frequency of this habitat)-

Evolution of B in a constant environment (sl = s2 = x = y)   From the previous sections it can be inferred that when the environment is neutral for both alleles A and a and when choice is not costly ($d = 1$), then both loci are neutral (i.e., $p' - p = 0$ and $q' - q = 0$, whatever p and q). Appendix C, however shows an analytical resolution of this point.

If in such a case choice is costly ($d < 1$), then B (choosy allele) is eliminated. In this situation, the adaptive aliele (A or a) the most associated with B, or choosing the rarest habitat (while associated with B), will be eliminated by hitch-hiking. This stays true as long as the linkage disequilibrium, between the two loci, or the frequency of B is not null.

EVOLUTION OF HETEROZYGOSUY AND SYMPATRIC SPECIATION

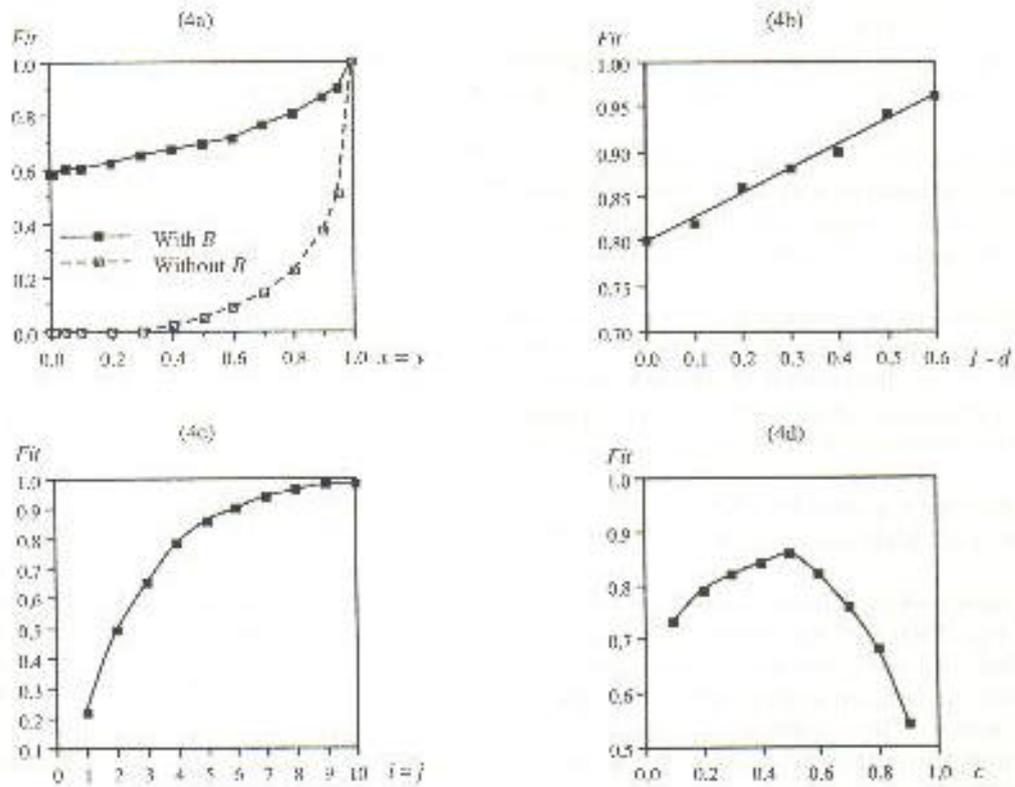From the previous sections it is known that in the hard selection model, for each locus, only

Figure 4. Heterozygote deficits (F_{it}) as a function of: (a) the selective costs encountered in the less favourable habitat (x and y); (b) the propagule death probability while moving from one patch to another ($1 - d$); (c) the number of allowed attempts (i *and* j) for the choosy genotypes; and (d) habitat 1 frequency (c). In (b), (c) and (d) B (choosy allele) is fixed. The values of each parameter, for each kind of simulation, are given in the text. Selection is soft.

fixation or neutrality can occur, but not stable polymorphism. For this reason the 'hard selection model' cannot lead to the maintenance of two genetic entities on two habitats, while ' the soft selection model does. Let us then consider, in the soft selection model, the influence' of each parameter on the F_{it} at equilibrium. F_{it} is the deficit of heterozygotes in the whole population and is here equivalent to an F_{st} ($F_{is} = 0$ because of local panmixia).

When there is no habitat selection (b fixed)

For the sake of simplicity, on an arbitrary basis, we study only the influence of the adaptation parameters x and y ($s1 = s2 = 0$, $c = 0.5$). As shown in Fig. 4(a), F_{it} increases exponentially with x and y, but reaches high values only for very high x and y. For instance, a value x = $y = 0.95$ is required for F_{it} to reach 0.51. This means that when the survival of individuals, which have settled in the wrong habitat, is as low as 0.05, then only 51% of the expected heterozygotes are missing. This suggests that only high levels of selection can be detected in natural populations submitted to such conditions.

When B iv fixed

In this section, and unless specified otherwise, it is arbitrarily considered that $x = y = 0.8$, $s1 = s2 = 0$, $d = 0.8$, $c = 0.5$ and $i = j = 5$, each term being varied one at a time.

Influence of the adaptation parameters x and y   Here it is considered that $d = I$ so that the results can be compared with the results of section 6(a). As shown in Fig. 4(a), when B is fixed, the value of F_{it} is much higher ($F_{it} = 0.58$ for $x = y = 0$), than when b is fixed (*Fit* $= 0$ for $x = y = 0$). Note that F_{it} reaches 0.86 when $x = y = 0.9$ (against only 0.37 when not choosing).

Infl ence of d (propagule dispersal survival)   Figure 4(b) shows an increase of F_{it}, when d decreases from 0.8 to 0.4 (for $d = 0.3$ B is selected against). Individuals which tend to reduce the F_{it}, are those which reach the ultimate try (here the fifth one) and settle in the first habitat they encounter. If mortality, at each attempt, increases, then the number of propagules surviving to the ultimate try (those which settle anywhere) decreases.

Influence of the number of tries (i and j) F_{it} increases asymptotically with i and j (Fig.4(c)) and can reach high values (0.98 for $i = j = 9$).

Influence of the habitat frequency c   A parabolic like relationship between F_{it} and c is observed in Fig. 4(d). The maximum is obtained for $c = 1/2$ (then $F_{it} = 0.86$). This is because, for low values of i and j, the more a given habitat becomes preponderant and the more the propagules reaching their last attempt will fall (and settle) in it, thus increasing the encounter between alleles A and a. The asymmetry of the curve (F_{it} decreases faster when c increases) is another consequence of the dominance of A. When c decreases, less heterozygotes (surviving better in habitat 1) are produced than when c increases.

Discussion

The novel aspect of the approach presented here is the choice of the most favourable habitat being epistatically determined by the adaptation locus. This allows the simultaneous study of the evolution of both polymorphism and habitat selection in a variable environment. This, to our knowledge, has never been formally dealt with. The main results obtained are the following.
The principle by which variability may not be maintained on a character governing habitat selection can be extended to both soft and hard selection models. This confirms the results of Doyle (1975), who studied the evolution of habitat selection alone, with respect to one habitat, in a hard selection framework.
   Local mating does not alter the conditions for polymorphism in soft selection models (Strobeck, 1974). We have been able to extend this notion to hard selection models and to models involving a genetically induced habitat selection.
This might appear to contradict the conclusions of some authors (e.g., Hoekstra *et al.,* 1985; Hedrick, 1986, 1990b) who use a 'Maynard-Smith (1966)' type of habitat selection. In Maynard-Smith's model, only females select for habitat. They tend to lay eggs in the habitat where they were born. Thus, in case of global panmixia, only the genes of females are involved in the process of habitat selection (genes of males may come from any habitat). But when mating occurs within each habitat, the eggs carried by the females, which are habitat selectors, contain genes from males and females, both involved in the habitat selection process. Here, local mating widens the conditions for protected polymorphism because it leads more genes to select for habitat.
   If adaptation depended on more than one locus, local mating, with habitat selection, would

certainly represent an advantage because it would allow for the conservation of the fittest allelic combinations. In the case of global random mating, the production of less fit combinations at each generation would enhance the evolution of true assortative mating. This kind of process would then appear very close to what Rosenzweig (1978) defines as competitive speciation.

Adaptation to a variable environment under soft selection conditions may, in its own right, maintain an adaptive polymorphism (Levene, 1953; Maynard-Smith, 1962; Prout, 1968; Gliddon and Strobeck, 1975), but requires very severe conditions (Maynard-Smith and Hoekstra, 1980).

Two mechanisms appear powerful enough to broaden these conditions. These are partial migrations between habitats (e.g. Moran, 1959; Eyland, 1971) and habitat selection (e.g. Maynard-Smith, 1966; Templeton and Rothman, 1981; Garcia-Dorado, 1987; Hedrick, 1990a).

In such works, habitat selection significantly broadens the conditions for protected polymorphism only when genetically determined by pleiotropy. This means that the same genes determine adaptation and optimal habitat selection. The evolution of habitat selection itself is not studied.

The epistatic assumption used in our models allows this study. Moreover, this keeps the genetic correlation needed between habitat selection and adaptation. Here, the Doyle's (1975) process of habitat selection provides a mechanistic and reasonably realistic description of the phenomenon. We have not found examples which would demonstrate unequivocally the relevance of such habitat selection mechanisms. Nevertheless, results obtained by several authors may be interpreted as such (Olson, 1985; Anstensrud and Schram, 1988; Emson and Mladenov, 1987).

When selection is soft, and if the choice aliele is selected for, polymorphism is always maintained at the adaptative locus. The important factor will therefore be the selection of this choice allele. In the soft selection model, density dependence does not allow the evolution of habitat preference in the absence of a previous adaptive polymorphism. In the absence of competition with other phenotypes, it is more favourable to exploit all resources, even the less favourable ones, than to exploit only one of these resources. The only advantage of our soft selection model, compared with that of Levene, is that this previous polymorphism does not necessarily need to be stable, but here the survival of the propagules during dispersal appears critical. This represents an important conclusion in as far as many papers, dealing with the maintenance of adaptive phenotypes, indeed involve habitat selection as a fixed character (but see Rosenzweig, 1987).

Under hard selection, unlike under soft selection, conditions for the maintenance of poly-morphism are not broadened by habitat selection, as it does not allow maintenance of any adaptive polymorphism. Habitat selection, however, does not require adaptive polymorphism in order to evolve, like in the soft selection model. It ensues, therefore, as postulated by Futuyma and Moreno (1988), that hard selection is more favourable to habitat specialization than soft selection.

Under hard selection, the impossibility of maintaining polymorphism of choice and adaptation could conceivably account for the existence of empty habitats. Once the habitat preference allele establishes itself, it does so with just one adaptation allele exploiting one of the two habitats. The other habitat, although available, remains empty. This could explain several observations concerning many phytophagous insects (e.g. Smiley, 1978; Myers *et al.,* 1981; Anderson *et al.,* 1989), and also some parasitic worms (e.g. Rhode, 1979), where the organisms studied appear not to exploit all the available habitats. It must be noted, however, that hard selection models are very unstable. A simple mutation or disturbance of the environment may bring everything into question.

Coexistence of several adaptive alleles is, thus, impossible in the context of hard selection. Sympatric speciation is therefore an evolutionary phenomenon incompatible with this type of situation.

In the soft selection model, on the other hand, if habitat preference is selected for, it may cause, within the population, very severe deficiencies in beterozygotes at the adaptive locus. These deficiencies will be all the more severe as the selective costs of being in the wrong niche, the maximum number of attempts (i, j) and the dispersal costs will be high. A similar availability of the different niches increases these deficiencies as well.

This must represent sympatric speciations. Indeed, from this starting point, selection will automatically tend to improve the specialization of each phenotype, that is, to improve the isolation. It is worth pointing out that habitat preference is not selected for to produce a genetic isolation but to better adapt individuals. Here, speciation occurs only as a by product of habitat selection with local mating.

Nevertheless, selection of habitat preference appears to follow fairly restrictive conditions within a soft selection framework. Maynard-Smith (1966) states that a stable polymorphism in a variable environment is a prerequisite for sympatric speciation. In our soft selection model, these conditions appear to be widened, but not very much, and only in the case of no cost to habitat selection.

The experiment carried out by Thoday and Gibson required a strong disruptive selection to result in genetic isolation, and has apparently proven impossible to repeat (e.g. Thoday and Gibson, 1970). The work of Rice and Salt (1988) gave rise to some spectacular results by selecting habitat preference. The selection used, however, was also very severe and had little in common with natural conditions. The natural examples (e.g. Tauber and Tauber, 1977; Gibbons, 1979; Wood and Guttman, 1983) can be explained by previous allopatric isolations and appear dubious to certain authors (e.g. Futuyma and Mayer, 1980). Nevertheless, the extreme diversity found in certain kinds of organisms such as parasites (e.g. Rhode, 1979) or mutualists (e.g. Kjellberg *et al.*, 1987) could appear as evidence in favour of this kind of speciation. Even in the fossil records, some data strongly suggest the importance of nonallopatric speciation in the evolutionary process (Rosenzweig and Taylor, 1980; Rosenzweig and McCord, 1991).

Alternatively, a model with some hard selection, allowing the emergence of the choice allele without a previous adaptive polymorphism, and with some soft selection, allowing the maintenance of adaptive polymorphisms with habitat selection, could provide a more convincing theoretical way to reach sympatric speciation.

Many studies have demonstrated discrimination of resources by organisms (reviews by Fox and Morrow, 1981; Hedrick, 1986, Futuyma and Moreno, 1988). It may be suspected that this has a genetic basis (Futuyma and Moreno, 1988). Nevertheless, it would appear that few probing examples of maintenance of polymorphism within a variable environment exist in nature (Futuyma and Mayer, 1980). Wasserman and Futuyma (1981) and Futuyma and Moreno (1988) also emphasize the lack of examples where resource preference is correlated with adaptation (at least in the case of phytophagous insects). In our soft selection model, the competition with a specialist provides a strong enough mechanism to explain habitat specialization of genotypes which do not survive better in any habitat. Nevertheless, several experimental or natural observations show some evidence of the maintenance of diversity with habitat selection (e.g. Tauber and Tauber, 1977; Christensen, 1977; Cavener, 1979; Jones and Probert, 1980; De Meeûs *et al.*, 1990).

On the other hand, the hard selection model can explain a change in the preferred resources from one locality to the other. In that model only one kind of resource can be used, and it is very sensitive to any changing in availability or quality of the habitat exploited.

In conclusion, it appears that soft and hard selection models show some radical differences. These differences may prove to be useful comparative tools for the description and interpretation of the evolution of natural populations. It is quite likely, however, that natural populations do

not correspond to either totally hard or soft selection models, but to intermediate situations. Nevertheless, the patterns observed in natural populations should allow an estimation of which mode of regulation dominates these populations (density dependence or otherwise). Specialized species, that refuse to exploit all the suitable habitats (available and empty), and are very sensitive to disturbances (genetic and/or environmental), must be regulated by hard selection (no density dependence). On the other hand, the regulation of species for which no available habitat is vacant and where these habitats are sympatrically exploited by different entities, will more likely correspond to a soft selection model (with density dependence).

Acknowledgements

References

Anderson, S. S., McRea, K. D., Abrahamson, W. G. and Hartzel, L. M. (1989) Host genotype choice by the ball gallmaker *Eurosta solidagitiis* (Diptera Tephritidae). *Ecology* **70,** 1048 54.

Anstensrud, M. and Schram, T. A. (1988) Host and site selection by the larval stages and adults of the parasitic copepod *Lernaeenicus sprattae* (Sowerby)(Copepoda, Pennellidae) in the Oslofjord. *Hydro-biologia* **167/168,** 587-95.

Barton, N. and Clark, A. (1990) Population structure and processes in evolution. *Population Biology: Ecological and Evolutionary Viewpoints* (K. W6hrmann and S. K. Jain, eds), pp. 115-73. Springer-Verlag, Berlin, Germany.

Bush, G. L. (1975) Sympatric speciation in phytophagous parasitic insects. *Evolutionary Strategies of Parasitic Insects and Mites* (P. W. Price, ed.), pp. 187-204. Plenum, New York, USA.

Cavener, D. (1979) Preference for cthanol in *Drosophila melanogaster* associated with the alcohol deshydrogenase polymorphism. *Behavior. Genet.* **9,** 359-65.

Christensen, B. (1977) Habitat preference among amylase genotypes in *Asellus aquaticus* (Isopoda, Crustacea. *Hereditas* **87,** 21-6.

De Meeûs, T., Renaud, F. and Gabrion, C. (1990) A model for studying isolation mechanisms in parasite populations: the genus *Lepeophtheirus* (Copepoda, Caligidae). *J. Exp. Zool.* **254,** 207-14.

Dempster, E. R. (1955) Maintenance of genetic heterogeneity. *Cold Spring Harb. Symp. Quant. Biol. Sci.* **20,** 25-32.

Diehl, S. R. and Bush, G. L. (1989) The role of habitat preference in adaptation and speciation. *Speciation and its consequences* (D. Otte and J. A. Endier, eds) pp. 345-85. Sinauer Associates, Sunderland, MA, USA.

Doyle, R. W. (1975) Settlement of planktonic larvae: a theory of habitat selection in varying environments. *Am. Nat.* **109,** 113-26.

Emson, R. H. and Mladenov, P. V. (1987) Britticstar host specificity and apparent host discrimination by the parasitic copepod *Ophiopsyllus red ctus. Parasitology* **94,** 7-15.

Eyland, E. A. (1971) Morail's island migration model. *Genetics* **69,** 399-403.

Felsenstein, J. (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* **35,** 124-38.

Fox, L. R. and Morrow, P. A. (1981) Specialization: species property or local phenomenon? *Science* **211,** 887-93.

Futuyma, D. J. and Mayer, G. C. (1980) Non-allopatric speciation in animals. *Syst. Zool.* **29,** 254-71.

Futuyma, D. J. and Moreno, G. (1988) The evolution of ecological specialization. *Ann. Rev. Ecol. Syst.* **19,** 207-33.

Garcia-Dorado, A. (1986) The effect of niche preference on polymorphism protection in a heterogeneous environment. *Evolution* **40**, 936-45.

Garcia-Dorado, A. (1987) Polymorphism from environmental heterogeneity: some features of genetically induced niche preference. *Theor. Pop. Biol.* **32**, 66-75.

Gibbons, J. R. H. (1979) A model for sympatric speciation in *Megarhyssa* (Hymenoptera, Ichneumonidae): competitive speciation. Am. *Nat.* **114**, 719-41.

Gliddon, C. and Strobeck, C. (1975) Necessary and sufficient conditions for multiniche polymorphism in haploids. *Am. Nat.* **109**, 233-5.

Hedrick, P. W. (1986) Genetic polymorphism in heterogeneous environment: a decade later. *Ann. Rev. Ecol. Syst.* **17**, 535-66.

Hedrick, P. W. (1990a) Genotypic habitat selection: a new model and its application. *Heredity* **65**, 145-9.

Hedrick, P. W. (1990b) Theoretical analysis of habitat selection and the maintenance of genetic variation. *Ecological and Evolutionary Genetics of Drosophila* (J. S. F. Baeker, W. T. Starmer and R. J. McIntyre, eds), pp. 209-27. Plenum, New York, USA.

Hoekstra, R. F., Bijlsma, R. and Dolman, A. J. (1985) Polymorphism from environmental heterogeneity: models are only robust if the heterozygote to the favoured homozygote in each environment. *Genet. Res. Cambr.* **45**, 299-314.

Kjellberg, G., Gouyon, P. H., Ibrahim, M., Raymond, M. and Vaideyron, G. (1987) The stability of the symbiosis between dioecious figs and their pollinators: a study of *Ficus carica* L. and *Blastophaga psenes L. Evolution* **41**, 693-704.

Jones, J. S. and Probert, R. F. (1980) Habitat selection maintains a deleterious allele in a heterogeneous environment. *Nature* **287**, 632-3.

Levene, H. (1953) Genetic equilibrium when more than one ecological niche is available. *Am. Nat.* **87**, 331-3.

Maynard-Smith, J. (1962) Disruptive selection, polymorphism and sympatric speciation. *Nature* 195, 60-2.

Maynard-Smith, J. (1966) Sympatric speciation. Am. *Nat.* **100**, 637-49.

Maynard-Smith, J. (1970) Genetic polymorphism in a varied environment. Am. *Nat.* 104, 487-90.

Maynard-Smith, J. and Hoekstra, R. F. (1980) Polymorphism in a varied environment: how robust are the models? *Genet. Res. Camb.* **35**, 45-57.

Mayr, E. (1963) *Populations, Species and Evolution.* Harvard University Press, Cambridge, MA, USA.

Mayr, E. (1982) Speciation and macroevolution. *Evolution* **36**, 1119-32.

Moran, P. A. P. (1959) The theory of some population genetics effects of population subdivision. *Aust. J. Biol. Sci.* **12**, 109-16.

Myers, J. H., Monro, J. and Neil, M. (1981) Egg clumping, host plant selection and population regulation *in Cactoblastis cactorum (Lepidoptera). Oecologia* **51**, 7-13.

Olson, R. 0. (1985) The consequences of short-distance larval dispersal in a sessile marine invertebrate. *Ecology* 66, 30-9.

Prout, T. (1968) Sufficient conditions for multiple niche polymorphism. *Am. Nat.* **102**, 493-6.

Rausher, M. D. (1984) The evolution of habitat preference in subdivided populations. *Evolution* **38**, 596-608.

Rhode, K. (1979) A critical evaluation of intrinsic and extrinsic factors responsible for niche restriction in parasites. *Am. Nat.* **114**, 648-71.

Rice, W. R. (1984) Disruptive selection on habitat preference and the evolution of reproductive isolation: a simulation study. *Evolution* **38**, 1251-60.

Rice, W. R. (1987) Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. *Evol. Ecol.* **1***,* 301-14.

Rice, W. R. and Salt, G. W. (1988) Speciation via disruptive selection: experimental evidence. *Am. Nat.* **131**, 911-17.

Rosenzweig, M. L. (1978) Competitive speciation. *Biol. J. Linn. Soc.* **10***,* 275-89.

Rosenzweig, M. L. (1987) Habitat selection as a source of biological diversity. *Evol. Ecol.* **1***,* 315-30.

Rosenzweig, M. L. and McCord, R. D. (1991) Incumbent replacement - evidence for long term evolutionary progress. *Paleobiology* **17**, 202-13.

Rosenzweig, M. L. and Taylor, J. A. (1980) Speciation diversity in Ordovician invertebrates: filing niches quickly and carefully. *Oikos* **35***,* 236-43.

Smiley, **J.** (1978) Plant chemistry and the evolution of specificity: new evidence from *Heliconius* and *Passiflora. Science* **201,** 745-7.

Strobeck, C. (1974) Sufficient conditions for polymorphism with N niches and M mating groups. *Am. Nat.* **108,** 152-6.

Tauber, C. A. and Tauber, M. J. (1977) A genetic model for sympatric speciation through habitat diversification and seasonal isolation. *Nature* **268,** 702-5.

Tauber, C. A. and Tauber, M. J. (1989) Sympatric speciation in insects: perception and perspective. *Speciation and Its Consequences* (D. Otte and **J.** A. Endier, eds), pp. 307-44. Sinauer Associates, Sunderland, MA, USA.

Templeton, A. R. and Rothman, E. D. (1981) Evolution in fine grained environments. 11. Habitat selection as a homeostatic mechanism. *Theor. Popul. Biol.* **19,** 326-40.

Thoday, **J.** M. and Gibson, **J.** B. (1970) The probability of isolation by disruptive selection. Am. *Nat.* **104,** 219-30.

Wallace, B. (1968) Polymorphism, population size, and genetic load, *Population Biology and Evolution* (R. C. Lewontin, ed.), pp. 87-108. Syracuse University Press, Syracuse, NY, USA.

Wallace, B. (1975) Hard and soft selection revisited. *Evolution* **29,** 465-73.

Wasserman, S. S. and Futuyma, D. J. (1981) Evolution of host plant utilisation in laboratory populations of the southern cowpea weevil, *Callosobruchus maculatus* Fabricius (Coleoptera, Bruchidae). *Evolu-tion* **35,** 605-17.

Wilson, D. S. and Turelli, M. (1986) Stable underdominance and the evolutionary invasion of empty niches. *Am. Nat.* **127,** 835 50.

Wood, T. K. and Guttman, S. 1. *(1983) Euchenopa binotata* complex: sympatric speciation? *Science* **220,** 310-12.

Appendixes

Throughout the appendixes, *i* and *j* (maximum number of attempts) are assumed to be large enough so that AB settle only in habitat I and aaB only in habitat 2. When an equilibrium is reached, linkage disequilibrium can be neglected.

$G(x)$ =initial frequency of genotype x, in the propagule pool;

$Gi(x)$ =genotypic frequency in habitat i (i = 1 or 2);

$G'(x)$ = genotypic frequency in the next generation propagule pool;

$p, p1, p2, p'$ = frequencies of allele A;

$q, q1, q2, q'$ = frequencies of allele B;

c = frequency of habitat 1;

d = dispersal survival when looking for a patch;

$s1, x, s2, y$ = within habitat survival costs of genotypes A and aa (see Table 1)

APPENDIX A

In this section, it is assumed that B (choosy allele) is fixed. The genotypic frequencies in each habitat after migration are thus:

in habitat 1:
$$G1(AA) = G(AA)cd[1-(1-c)d]^{-1}(W1)^{-1}$$
$$G1(AA) = G(Aa)cd[1-(1-c)d]^{-1}(W1)^{-1}$$
$$G1(aa) \quad 0$$

with $W1 = [1-G(aa)]cd[1-(1-c)d]^{-1}$

this gives:
$$G1(AA) = G(AA)[1-G(aa)]^{-1}$$
$$G1(Aa) = G(Aa)[1-G(aa)]^{-1}$$

in habitat 2:
$$G2(AA) = G2(Aa) = 0 \text{ and } G2(aa) = 1$$

We thus have the following gene frequency of A in each habitat:
$$p1 = G1(AA) + 1/2G1(Aa) = p[1-G(aa)]^{-1} \text{ and } p2 = 0$$

Local reproduction does not affect local gene frequencies. Thus p', the frequency of A in the next generation, can be calculated for both soft and hard selection models:

(a)   In the soft selection model

$$p' = cp1 + (1\text{-}c)p2 = cp1 = cp[\text{l-}G(\text{aa})]^{-1}$$

At equilibrium $p'\text{-}p = 0$ or $G(\text{aa}) = 1\text{-}c$. Thus, at equilibrium, $p1 = p[\text{l-}G(\text{aa})]^{-1}$ = p/c. This means that the equilibrium frequency of A is equal to c, the frequency of its own habitat. It is easily seen that this equilibrium is stable.

(b) In the hard selection model
At generation $t + 1$ the new frequency of A in the propagule pool is

$$p' = \{W1p1 + W2p2\} \{W1 + W2\}^{-1} = p(1\text{-}sl)c[\text{l-}(1\text{-}c)d]^{-1} \{W1 + W2\}^{-1}$$

The evolution of allele A is given by $p'\text{-}p$, the sign of which depends only on the sign of $(1\text{-}sl)c[\text{l-}(1\text{-}c)d]^{-1} - (1\text{-}s2)(1\text{-}c)[1\text{-}cd]^{-1}$, which is independent of the allelic or genotypic frequencies.

APPENDIX B

In this section it is considered that A is fixed at the adaptive locus. Given that A is fixed, genotypes BB and Bb choose exclusively habitat 1. Genotypes bb settle in habitat I with a probability c and in habitat 2 with a probability I -c. Then, after colonization, the local genotypic frequencies are

in habitat 1 $G1(\text{BB}) = G(\text{BB})cd[\text{l-}(\text{l-}c)d]^{-1}(W1)^{-1}$
             $G1(\text{Bb}) = G(\text{Bb})cd[\text{l-}(\text{l-}c)d]^{-1}(W1)^{-1}$
             $G1(\text{bb}) = G(\text{bb})cd(W1)^{-1}$

with $W1 = cd\{ [1\text{-}G(\text{bb})] [1\text{-}(\text{l-}c)d]^{-1} + G(\text{bb})\}$

or

             $G1(\text{BB}) = G(\text{BB}) [1 - G(\text{bb}) (1\text{-}c)d]^{-1}$
             $GI(\text{BB}) = G(\text{Bb}) [1 - G(\text{bb}) (1\text{-}c)d]^{-1}$
             $Gl(\text{bb}) = G(\text{bb}) [1\text{-}(\text{l-}c)d] [1 - G(\text{bb}) (1\text{-}c)d]^{-1}$
in habitat 2 $G2(\text{BB}) = G2(\text{Bb}) = 0$, $G2(\text{bb}) = 1$ and $W2 = (1\text{-}c)dG(\text{bb})$

Thus the local gene frequencies of allele B are:
        $q1 = GI(\text{BB}) + 1/2Gl(\text{Bb}) = q[\text{l-}G(\text{bb}) (1\text{-}c)d]^{-1}$ and $q2 = 0$.

Thus q', the overall frequency of B in the next generation, can be calculated for soft and hard selection models.

In the soft selection model
        $$q' = cq1 + (1\text{-}c)q2 = cq[1 - G(\text{bb}) (1\text{-}c)d]^{-1}$$

Then $q'\text{-}q = q[c - 1 + G(\text{bb}) (1\text{-}c)d] [1 - G(\text{bb}) (1\text{-}c)d]^{-1}$
which has the sign of $-(1 - c) [1 - G(\text{bb})d]$ and is always negative. Thus b (no habitat preference) goes to fixation.

In the hard selection model
At generation $t + 1$, the frequency of allele B, in the new propagule pool is

$$q' = [W1q1] + W2q2] [W1 + W2]^{-1} = qc(1\text{-}sl) [1\text{-}(\text{l-}c)d]^{-1} [Wl + W2]^{-1}$$

Here $q'-q$ has the same sign as $(1-s1)cd[1-(1-c)d]^{-1} - (1-x)$. Similarly, when allele a is initially fixed, the evolution of the choice locus (B/b) is described by the sign of $(1-s2)(1-c)d[1-cd]^{-1} - (1-y)$.

APPENDIX C

It is now assumed that AA, Aa and aa have equal fitness in each habitat. The evolution of genotypic frequencies at both adaptive and choice loci are studied. Local mating does not change the alielic frequencies. The initial gene frequencies p and q of alleles A and B are:

$p = G(AABB)+G(AABb)+G(AAbb) + 1/2 [G(AaBB)+G(AaBb)+G(Aabb)]$

or, if we assume linkage equilibrium,

$p = G(AA) + 1/2G(Aa)$

Similarly,

$q = G(BB) + 1/2G(Bb)$

Now let migration and gamete production occur within each habitat. Local gene frequencies of A are then

in habitat 1

$p1 = G1(AABB)+G1(AABb)+G1(AAbb) + 1/2 [G1(AaBB)+G1(AaBb)+G1(Aabb)]$

or

$p1 = \{G(AABB)+G(AABb) + c[G(AAbb)+1/2G(Aabb)] + 1/2[G(AaBB)+G(AaBb)]\}(W1)^{-1}$ with

$W1 = G(AABB)+G(AABb)+G(AaBB)+G(AaBb) + c[G(AAbb)+G(Aabb)+G(aaBB)]$

If initial linkage equilibrium is assumed:

$G(AABB) = G(AA) G(BB)$ and $W1 = [1-G(bb)]$ fl-$G(aa)] + cG(bb)$

$p1 = p[I-G(bb) (1-c)] \{ 1 -G(aa) - G(bb) [1-c - G(aa)]\}^{-1}$

The frequency of A in habitat 2 is

$p2 = (1-c) [G(AAbb) + 1/2G(Aabb)] (W2)^{-1}$

with

$W2 = G(aaBB)+G(aaBb) + (1-c) [G(AAbb)+G(Aabb)+G(aabb)]$

It is of note that $W2 = 1 - W1$.

Given initial linkage equilibrium, we have

$W2 = G(aa) [1-G(bb)] + (1-c)G(bb) = 1-W1$

and

$p2 = p(1-c)G(bb) \{ G(aa) - G(bb) [1-c - G(aa)]\}^{-1}$

Similarly, the gene frequency of B is

in habitat I

$q1 = \{G(AABB)+G(AaBB) + 1/2 [G(AABb)+G(AaBb)]\}^{-1} (W1)^{-1}$

and in habitat 2

$q2 = \{ G(aaBB) + 1/2G(aaBb) \} (1-Wl)^{-1}$

with $Wlq1 + (1-Wl)q2 = q$.

In the soft selection model

In the next generation, the overall gene frequency of A is thus

$p' = cp1 + (1-c)p2$, which gives

$p' = p\{c[1 - (1-c)G(bb)] (Wl)^{-1} + (1-c)^2 G(bb) (1-Wl)^{-1} \}$

The overall gene frequency of B is

$q' = cq1 + (1-c)q2$

Thus, the general evolution of B may be given by

$q'-q = cq1 + (1-c)q2 - Wlq1 - (1-Wl)q2 (Wl-c) (q2-q1)$

Given linkage equilibrium we have then

$q' = q\{c[1-G(aa)](W1)^{-1} + (1-c) G(aa) (1-W1)^{-1}\}$

At equilibrium, *p'=p and q'=q* and, therefore,

(1) $c[1-(1-c) G(bb)] (W1)^{-1} + (1-c)^2 G(bb) (1-Wl)^{-1} = 1$

and

(2) $c[1-G(aa)](Wl)^{-1} + (1-c) G(aa) (1-W1)^{-1} = 1$

We can put Equation 1 and Equation 2 within the same equation:

$c[(G(aa) - (1-c)G(bb)] (W1)^{-1} = (1-c) [(1-c)G(bb) - G(aa)] (W2)^{-1}$

The case when $G(aa)\neq(1-c) G(bb)$ is considered first. Here, the equilibrium conditions write:

$c(1-Wl) = (1-c) Wl$, or $Wl = c$

Thus

(3) $[1 - G(bb)] [1-G(aa)] + cG(bb) = c$

Fixation of b ($G(bb) = 1$) is a solution of Equation 3.

If $G(bb)\neq1$, the solution is then

(4) $G(aa) = 1-c$

This means that there are as many genotypes aa as the amount of habitat chosen by phenotype [aaB]. When Equation 4 is verified, $q' = q$ for all values of q (neglecting hitch-hiking due to linkage disequilibrium), and B is neutral.

If $G(bb) = 1$, b is fixed, and $p' = p$ for all values of *p,* as A is then neutral.

The second situation considered is the case when $G(bb) (1-c) = G(aa)$. It can be shown that, in this case, the only solution to equilibrium is $G(bb) = 1$ (and $G(aa) = 1-c$), that is fixation of b, and thus neutrality of A/a.

For a haploid population when selection is soft

Here we consider a haploid population with thus only four genotypes: AB, Ab, aB and ab. The environment is assumed neutral. The initial allelic frequencies are

$p = G(AB) + G(Ab)$ for A and $q = G(AB) + G(aB)$ for B.

After settlement genotypic frequencies are
in habitat 1 $Gl(AB), G(AB) (Wl)^{-1}, Gl(Ab) = G(Ab)c(Wl)^{-1}$
            $GI(AB) = 0, G1(ab) = G(ab)c(Wl)^{-1}$
where $Wl = G(AB) + c[G(Ab) + G(ab)] = G(AB) + c(l-q)$
in habitat 2 $G2(AB) = 0, G2(Ab) = G2(Ab) (1-c) (W2)^{-1}$
            $G2(aB) = G(aB) (W2)^{-1}, G2(ab) = G(ab) (1-c) (W2)^{-1}$
where $W2 = G(aB) + (1 - c) (1 - q) = 1 - Wl$.
The frequencies of A and B are then

$$p1 = [G(AB) + cG(Ab)] (Wl)^{-1} , \quad ql = G(AB) (Wl)^{-1}$$

in habitat 1 and

$$p2 = (1-c)G(ab)(1-Wl)^{-1}, q2 = G(aB) (1-Wl)^{-1}$$
in habitat 2. We thus note that Wlql + *(1-Wl)q2 q*In the next generation these frequencies are then

$$p' = cp1 + (1-c)p2 \text{ for A}$$

and

$$q' = cql + (1 -c)q2 \text{ for B}$$

The evolution of the frequency of B is described by

$$q'-q = cql + (1-c)q2 - Wlql - W2q2 = (Wl-c) (q2-ql)$$

$q'-q > 0$ means that

       (i) $Wl > c$ and $q2 < ql$

or

       (ii) $Wl < c$ and $q2 > ql$

that is, for inequalities i

     $G(AB) + c(1-q) > c$
and
     $G(AB) [G(AB) + c(1-q)]^{-1} < G(aB) [G(aB) + (1-c)(1-q)]^{-1}$

which is equivalent to

     $G(AB) > cq$ and $G(AB) < cq$

which is impossible.

It is easy to see that inequality ii leads to the same result. This means that $q'-q$ is always negative or null. Thus, B.is always selected against or neutral. It is also easy to demonstrate that $Wl = c$ and $q2 = ql$ both lead to the same condition: $G(AB) = cq$.. In that case, and if we neglect linkage disequilibrium, it means that $p = c$.

In the hard selection model

In the next generation, we know that $p' = W1\,p1 + W2\,p2$ and $q' = W1\,q1 + W2\,q2$. It is easily seen that $p' = p$ and $q' = q$, which means that there is no evolution at these loci (neutrality).