

Interaction of selection and recombination in the fixation of negative-epistatic genes

YANNIS MICHALAKIS¹* AND MONTGOMERY SLATKIN²

¹ Institut d'Ecologie, Université Pierre et Marie Curie, Bat. A 7^{ème} ét. CC 237, 7 quai St Bernard, 75252 Paris CEDEX 05, France.

E-mail: imichala@snv.jussieu.fr

² Department of Integrative Biology, University of California, Berkeley, CA 94720-3140, USA. E-mail: monty@mws4.biol.berkeley.edu

(Received 7 April 1995 and in revised form 20 July 1995)

Summary

We investigated the interaction of recombination and selection on the process of fixation of two linked loci with epistatic interactions in fitness. We consider both the probability of fixation of newly arising mutants (the static model) and the time to fixation under continued mutation (the dynamic model). Our results show that the fixation of a new advantageous combination is facilitated by higher fitness of the advantageous genotype and by weaker selection against the intermediate deleterious genotypes. Fixation occurs more rapidly when the recombination rates are small, except when selection against intermediate genotypes is weak and selection in favour of the double mutant is very strong. In these cases fixation is more rapid when the recombinant rate is large. Mutations of strong effects, deleterious when alone but beneficial when coupled, are fixed more easily than mutations of intermediate effects, at least for large recombination rates. Among the possible pathways the process of fixation might follow, independent substitutions lead to the fixation of the double mutant only when selection is weak. The relative importance of the other pathways depends on the interaction between recombination and selection. The coupled-gamete pathway (i.e. when the population waits until the double mutant appears and then drives it to fixation) is more important as selection intensity increases and the recombination rate is reduced. For all recombination rates, asymmetries in fitness of the intermediate genotypes increase the rate at which fixations occur. Finally, throughout the fixation process, the population will be monomorphic at least at one of the two loci for most of the time, which implies that there would be little opportunity to detect the presence of negative epistasis even if it were important for occasional evolutionary transitions.

1. Introduction

One view on the evolution of adaptive characters is that natural selection acts on genes largely independently of one another (Williams, 1992). The main advocate of this view was Fisher (1930), who argued that the fate of a gene depends on its average effect on fitness across all possible gene combinations, so that additive effects drive adaptive evolution. Negative epistatic gene interactions on fitness were proposed by Wright as an additional factor to be considered (Wright, 1931). According to Wright, a gene favourable in one given genomic context is likely to be unfavourable in other genomic contexts. Under Wright's view of adaptive evolution, epistatic interactions would create several 'adaptive peaks' in

average fitness, separated by 'valleys' representing genetic combinations of lower average fitness. Forces other than natural selection would be necessary to move a population from one adaptive peak to another.

The shifting-balance theory has recently been examined in several theoretical studies. Most studies have concentrated on the third phase, the spread of the advantageous genotype in a subdivided population once fixed in a single deme (Crow, Engels & Denniston, 1990, Barton, 1992, Kondrashov, 1992, Phillips, 1993). Several studies have considered the entire process but with emphasis on the third phase. For example Moore & Tonsor (1994) were concerned with the range of migration rates that would maximize the probability of peak shifts. They used a two-locus diploid model and showed that shifts happen for intermediate migration rates ($0.001 < m < 0.01$). At lower migration rates, shifts occur relatively frequently but the

* Corresponding author.

new genotype combination is not exported sufficiently often. For larger migration rates, fixation of the new genotype within demes is extremely rare. Because Moore and Tonsor were interested primarily in the effects of migration they used only one fitness matrix (with weak selection against intermediate genotypes and moderately strong selection in favour of the new genotype) and only one recombination rate.

Barton & Rouhani (1993), using a model of a polygenic character under disruptive selection as well as a single locus model with underdominance, showed that the shifting-balance process allows adaptation most efficiently if the number of migrants between demes lies just below a critical value, $N_m \approx 1$.

In this paper we concentrate on the initial stage of the shifting balance process, namely on what happens within a single deme. We use a two-locus two-alleles haploid model with negative epistatic interactions on fitness between loci (see the Model section for details). We have four main goals. First, to investigate the interaction of selection and recombination between loci and the way it affects the probability of fixation and the time to fixation of a new advantageous mutant. Particular aspects of this problem have been addressed by a few authors in the past. Kimura (1985) looked at the process of fixation of compensatory neutral mutations, while Takahata (1982) addressed the question of whether recombination facilitates or impedes the fixation of new advantageous genotypes. Rutledge (1970) arrived at a numerical solution of the probability of fixation of the advantageous mutant for the case where intermediate genotypes are almost lethal and the advantageous combination is initially present in the population. Our scope is more general in that we study a much larger set of interactions between selection, recombination and the initial situation of the population.

Second, we want to evaluate the relative importance of the various pathways that could lead to the fixation of a new advantageous genotype. Several pathways to fixation are possible: (a) one of the two loci could first be fixed for a mutant, thereby 'resting' in the adaptive valley while awaiting the appearance of the mutation at the other locus (the 'independent substitutions' path), (b) both loci could be polymorphic so that the advantageous genotype is produced by recombination (the 'recombination' path), or (c) the advantageous genotype could be produced by mutation and then driven to fixation (the 'coupled gamete' path).

Third, we address the question of whether mutations with intermediate effects, deleterious when alone but beneficial when coupled, will be fixed more easily than mutations with strong effects.

Finally, we find the proportion of time of the fixation process during which the population is polymorphic at both loci. This quantity is important because experimental investigations will be able to show that variability in fitness is due to epistatic interactions (and not additive effects), only if genetic

variability is present at both loci. Otherwise, an experimental analysis would indicate that genetic variability for fitness is attributable to additive effects.

We first ignore mutation, and consider the recombination-selection interaction for fixed initial conditions (the static model). Of particular interest are the differences between the cases where the mutants are initially on a single gamete (i.e. in coupling) or initially on two gametes (i.e. in repulsion). These results help us interpret the results of the more realistic model (the dynamic model) where the population evolves under the joint effects of recombination, selection and mutation. We will show that much of the behaviour of the static model can be predicted from an analysis of what happens after one generation of selection and recombination.

Phillips (1996) used a numerical solution to a two-dimensional diffusion equation, a Fisher–Wright Markov approach and stochastic simulations to estimate the probability and time to fixation of a new genotypic combination. The cases he examined correspond to our static and dynamic models. He considered diploid models, which allowed him to take into account the effects of dominance in relative fitness (his 'metabolic pathway' model). Our simulation results are in agreement with Phillips' analytic and simulation results for comparable parameter values.

2. The model

We consider a haploid population of N individuals with discrete generations. We assume that fitness is controlled by two loci A and B with two alleles each. The two loci recombine at a rate r . There are four possible gamete types, AB , aB , Ab and ab , with respective frequencies x_1 , x_2 , x_3 and x_4 and relative fitnesses 1 , $1 - (1 + \alpha)s$, $1 - (1 - \alpha)s$ and $1 + t$ respectively. The parameter α indicates the degree of asymmetry ($\alpha \geq 0$; when $\alpha = 0$ the two intermediate genotypes have the same fitness $1 - s$; for values of s and α such that the fitness of a genotype would be negative, the fitness is set to 0), and t is the relative advantage of genotype ab which represents the new advantageous combination. Initially we will assume that $\alpha = 0$ and that $t = s$, so the single parameter s characterizes the selection regime. We will also consider cases in which both restrictions on selection are relaxed.

We consider two situations. In the first, only one copy of each mutant is present in the population (Phillips (1996) studied in much more detail the effect of the initial frequency of the mutants). In that case, we are interested in the probability (P) and time to fixation (T) of the advantageous genotype, ab , in the absence of mutation. The mutant alleles may initially be present in the same gamete, ab , or in two gametes Ab and aB . We call this situation the static case. Because we sometimes use strong selection coefficients,

the order of events in a generation is important. We consider both recombination before selection and the reverse. Recombination before selection corresponds to the situation in which selection acts after the formation of the gametes, so selection operates between the offspring of all individuals present in the previous generation. Selection before recombination corresponds to the situation in which selection acts before the formation of gametes, so selection affects the probability that each genotype reproduces. In other words, in the latter case selection tends to eliminate deleterious genotypes before they reproduce while in the former case after they have reproduced.

With recombination before selection (RS), the expected gamete frequencies are

$$x'_i = (x_i - \delta r D) w_i / W \tag{1}$$

where

$$\delta = 1 \text{ for } i = 1 \text{ and } 4 \text{ and } \delta = -1 \text{ for } i = 2 \text{ and } 3$$

$$D = x_1 x_4 - x_2 x_3$$

$$W = \sum (x_i - \delta r D) w_i$$

the prime denotes the frequencies in the succeeding generation, and w_i is the relative fitness of genotype i .

When selection acts before recombination (SR), we have

$$x'_i = x_{is} - \delta r D_s \tag{2}$$

with

$$x_{is} = x_i w_i / W$$

$$W = \sum x_i w_i$$

$$D_s = x_{1s} x_{4s} - x_{2s} x_{3s}$$

In our computer program selection and recombination are deterministic. They affect gametic (or adult) frequencies, and these frequencies serve as probabilities to randomly draw offspring to form the next generation. We checked our program by running simulations with no recombination, where the model behaves as a one-locus-four-alleles model, and compared the results to the results of diffusion approximations (Crow & Kimura, 1970). The results always agreed to the fourth decimal (results not shown).

The second model we consider is the dynamic model. In this case the population is initially fixed for AB . Genetic variability is introduced by mutation. We consider only one-way mutation at a rate μ from capital alleles to small alleles. Mutation is also considered as a deterministic process affecting gamete frequencies in a way analogous to recombination and selection as described in the previous paragraph.

In this case the fixation of the advantageous combination is certain and we were interested only in the time to fixation (T). In the dynamic model also, we considered both selection acting before recombination and after recombination. We checked this model by

comparing results we obtained with $n = 50$, $r = 0.5$, $\mu = 0.01$ and 21 values of $s = t$ (ranging from 0 to 1 at 0.05 intervals) with results Phillips obtained using his simulation program and found the results of the two programs to be in agreement.

3. Predictions for the static model in the first generation

Recombination and selection will have different effects depending on the initial conditions. When a coupled gamete is present in the population, recombination will act against it, breaking it down, while selection will act in its favour. On the other hand when the mutants are initially present as repulsion gametes we can divide the process in two phases. The first phase consists of the formation of at least one ab gamete. During this phase recombination acts in favour of the double mutant by creating it, while selection acts against it. During the second phase the roles of the two mechanisms are reversed.

We will show that what happens in the first generation is very important for the overall behaviour of the model. The quantity of interest is the expectation of the frequency of ab after the first generation.

3.1. ab present initially

In this case, the initial frequencies of the four gametes are $(N-1)/N$, 0, 0 and $1/N$ respectively. Because W is approximately 1, we have

$$x'_4 = \left(\frac{1}{N} - r \frac{N-1}{N^2} \right) (1+t) = \frac{1}{N} \left(1 - r \frac{N-1}{N} \right) (1+t) \tag{3}$$

for both the RS and the SR models. Thus, we expect both models to behave the same way under this initial condition. Furthermore, the fitnesses of the intermediate genotypes are not important, so we expect asymmetries in their fitnesses, as measured by α , to also not be important. On the other hand we expect the fitness of ab , $1+t$, to have a strong effect on P and T .

3.2. aB and Ab present initially

In this case, the initial frequencies of the four gametes are $(N-2)/N$, $1/N$, $1/N$ and 0 respectively, and we have

$$x'_4 = r \frac{1}{N^2} (1+t) \tag{4}$$

for the RS model, and

$$x'_4 = r \frac{1}{N^2} ((1-s)^2 - (s\alpha)^2) \tag{5}$$

in the SR model.

From these results, we predict that asymmetries in fitness do not have a strong effect in the RS model when the population initially contains only repulsion gametes because the process will be dominated by the fitness of ab .

Under the SR model, asymmetries are expected to influence P . The larger the asymmetry, the lower x_4' is. In the extreme case, when one of the intermediate genotypes is lethal P will be 0, because one of the two loci will be immediately fixed before recombination may create the double mutant. Because the effect of asymmetries depends also on the square of the value of s (see eqn. 5), we expect asymmetries to have an appreciable effect only when selection is strong.

In the SR model the fitness of the double mutant will affect P as well but for a different reason. In the RS case it directly influences the probability that the double mutant exists after the first generation (eqn. 4). In the SR model it will affect the probability of fixation of the double mutant, but only during the second phase of the process, i.e. after it was created by recombination.

Obviously what happens in the first generation does not tell the complete story. For example, after recombination has created the double mutant the population will behave as in the coupled gamete case. We expect the interaction between what happens in the first generation and the subsequent generations to lead to the existence of an intermediate recombination rate maximizing P .

We can summarize our predictions from the first generation analysis. When the two mutant alleles are introduced as a coupled gamete we expect the RS and SR models to behave the same way: P is expected to decrease monotonically with r and to increase monotonically with s . We do not expect asymmetries in fitness of the intermediate genotypes to have any effect, while the fitness of the double mutant is expected to have a large effect.

When the two mutant alleles are introduced as two repulsion gametes we expect the two models to behave similarly with respect to recombination but very differently with respect to selection. In both models some recombination is necessary for the formation of the double mutant, while recombination is detrimental once the double mutant is formed. Therefore under both models for a given selection coefficient we expect to find an intermediate recombination rate maximizing P .

The two models are expected to behave differently under changes in selection intensity, although in both models P is expected to increase monotonically with the intensity of selection on ab (t). Selection against the single mutants (s) is less important for the RS model and consequently asymmetry in the fitness of single mutants (α) would be expected to be unimportant. In contrast, the values of s and α are both important in the SR model because selection on the single mutants has more of an opportunity to act.

4. Results

4.1. Static model

Probabilities and times to fixation were estimated by running 20000 simulations for each parameter value combination when the initial state was one coupled gamete and 10^6 simulations for each parameter value combination when the initial state were two repulsion gametes. Fewer simulations were run for the coupled gamete initial condition because fixation probabilities were much higher.

(a) Mutant introduced as a single coupled gamete

Under both the SR and RS models the probability of fixation of ab increased monotonically with the selection coefficient and with the inverse of the recombination rate (Fig. 1). In both cases the time to fixation decreased as selection intensity increased, while it was almost insensitive to r (Fig. 2). Quantitatively the results of the two models were very close, the probability of fixation being slightly larger and the time to fixation slightly shorter for the SR model for large recombination rates, as predicted by the first-generation analysis. The selective advantage of the double mutant had a very strong effect for both models (Table 1 shows the results for the SR model; the RS model behaved in an exactly analogous way). As predicted, P increased and T decreased as t , the selective advantage of ab , increased. Asymmetries in the selection coefficients of the two intermediate genotypes (i.e. $\alpha > 0$) did not affect those quantities even for very strong selection coefficients. In fact, with tight linkage the fitness of the intermediate genotypes had almost no effect. For example in the SR model for $r = 0.02$ and $s = t = 0.60$, $P = 0.63$, while for $s = 0.30$ and $t = 2s = 0.60$, $P = 0.62$. These results agree with the predictions in the previous section.

The probability of fixation of the intermediate genotypes rapidly decreased with s and reached 0 for $s > 0.1$ (detailed results not shown).

(i) Effect of population size

The results were qualitatively the same for $N = 25$ and $N = 100$ (results not shown). The only difference was that for $N = 100$ and large r (i.e. $r = 0.5$) P no longer increased monotonically with the selection coefficient, but instead reached a minimum for intermediate values ($s = t = 0.20$). This behaviour was not predicted by the first generation analysis.

If the other parameters were unchanged, increasing N always decreased P . The difference was larger as r increased and the strength of selection decreased: in the absence of recombination P was always of the same order of magnitude for all values of N , for $r = 0.25$ differences in the order of magnitude of P appeared for values of $s = t = 0.20$, while for $r = 0.50$ such differences appeared for $s = t = 0.70$.

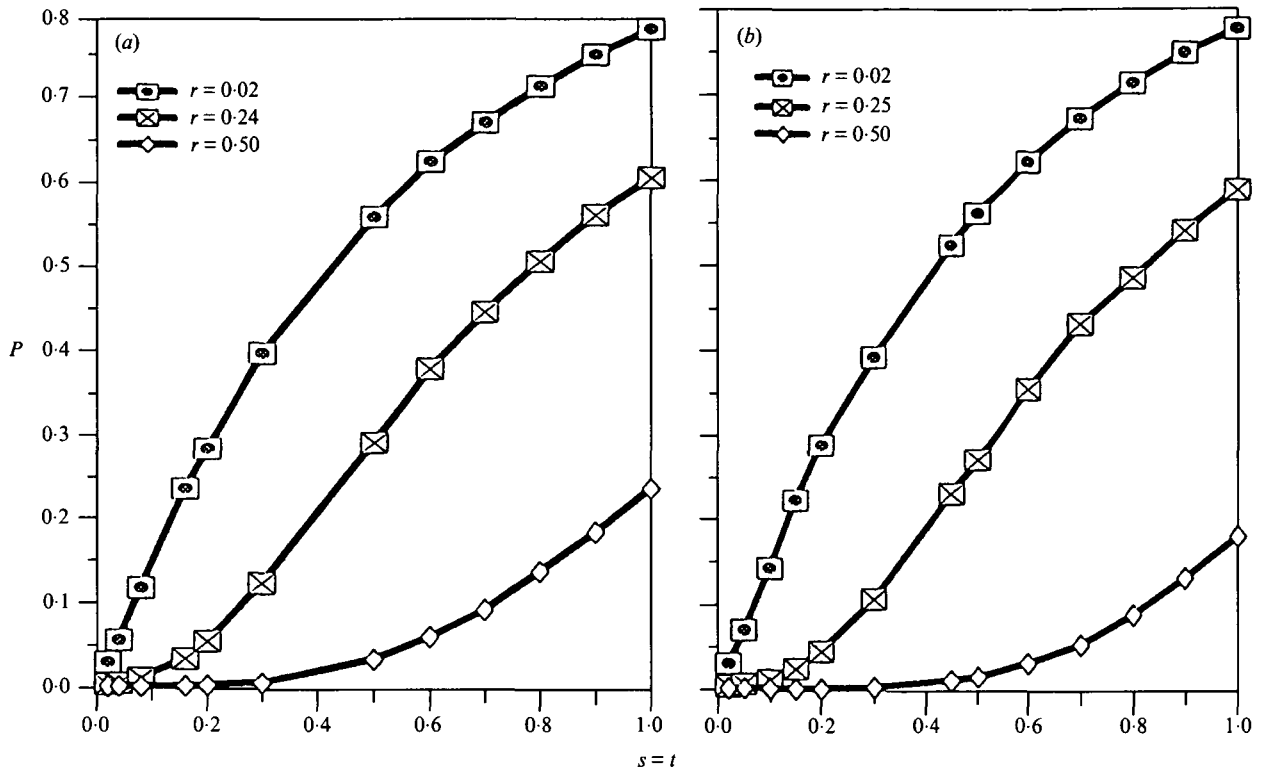


Fig. 1. Probability of fixation of ab when it is initially present in one copy. (a) SR model, (b) RS model. In all figures $N = 50$.

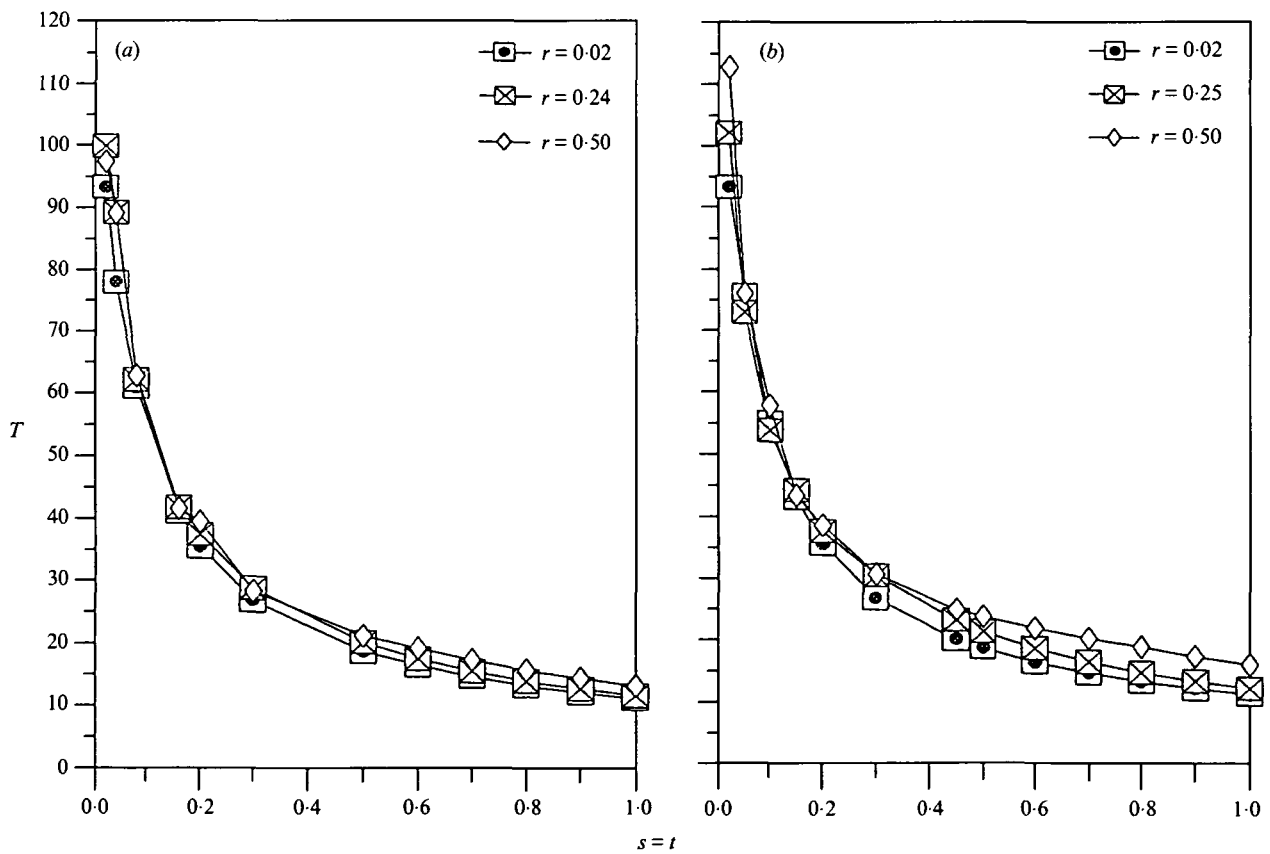


Fig. 2. Time to fixation of ab when it is initially present in one copy. (a) SR model, (b) RS model.

(b) Mutant introduced as two repulsion gametes

In this case, the order of events had a great influence on the results. Under the SR model P is not monotonic with r or s (Fig. 3a). We can recognize three domains of the parameter space: the small s domain, the small r domain, and the remaining domain. In the small s domain the situation is quasi-neutral. P increased with the recombination rate because the probability

that the double mutant is generated increases with r (see eqn. 5). In the small r domain, P decreased with s , because the probability that the double mutant is generated is very small, and increasing s increases the rate at which the single mutants are removed from the population. In the remaining domain recombination and selection interact to lead to intermediate rates maximizing P . The value of r maximizing P increased as s increased (Fig. 3a).

Table 1. Effect of the selective value of the ab ($1+t$) on the probability and time to fixation of ab

P_{ab}		$\alpha = 0$		
r	s	$t = s/2$	$t = s$	$t = 2s$
0.02	0.02	0.02080	0.03035	0.05705
	0.30	0.22580	0.39890	0.62170
	1.00	0.55210	0.78660	0.93305
0.24	0.02	0.00280	0.00395	0.00895
	0.30	0.02455	0.12415	0.37395
	1.00	0.29875	0.61165	0.85735
0.50	0.02	0.00150	0.00170	0.00265
	0.30	0.00050	0.00690	0.09610
	1.00	0.02655	0.24155	0.64325

T_{ab}		$\alpha = 0$		
r	s	$t = s/2$	$t = s$	$t = 2s$
0.02	0.02	100.3	93.3	84.2
	0.30	42.2	27.1	17.1
	1.00	19.0	12.1	8.1
0.24	0.02	117.0	99.9	95.6
	0.30	38.8	29.0	18.9
	1.00	19.5	12.6	8.0
0.50	0.02	108.0	97.5	100.7
	0.30	28.6	28.6	21.7
	1.00	20.0	14.2	8.5

Initial state: coupled gamete, SR model, $N = 50$.

The RS model was different from the SR model under changes in selection intensity but not under changes in the recombination rate (Fig. 3b). Indeed, for each value of s an intermediate value of r maximized P , and that value of r was approximately equal to that for the SR model. However, in contrast to the SR model, P increased as s increased in general. This increase was monotonic for small to intermediate recombination rates ($r < 0.25$) while P first decreased and then increased for larger r . This initial decrease for large recombination rates can be explained by the fact that s represents both the disadvantage of the intermediate genotypes and the advantage of ab . When r is large, ab is generated relatively easily but its fitness is not large enough to lead it to fixation while the large recombination rate tends to destroy it. As s increases the frequency of ab increases by selection at a rate faster than the rate that it is decreased by recombination, thus leading to an overall increase of P with s .

These results are all consistent with the predictions based on the analysis of the first generation, namely that the SR and RS models would behave similarly with respect to recombination but very differently with respect to selection, and in particular the decrease with the strength of selection in the SR model and the increase with the strength of selection in the RS

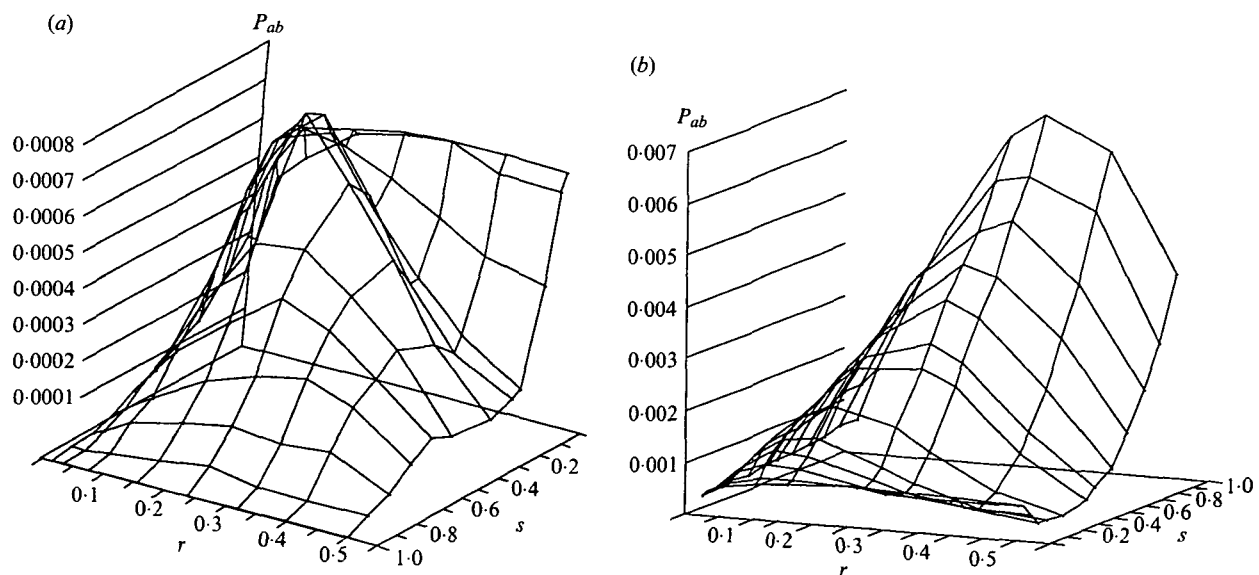


Fig. 3. Probability of fixation of ab when the initial population contains one copy of each repulsion gamete. (a) Shows results for the SR model and (b) for the RS model. Notice that the direction of the s axis is reversed between the two graphs, and that there is an order of magnitude difference between the probabilities of fixation obtained under the two models.

Table 2. Effect of the selective value of the *ab* ($1+t$) on the probability of fixation of *ab*

RS model				
<i>r</i>	<i>s</i>	$t = s/2$	$t = s$	$t = 2s$
0.02	0.02	0.000250	0.000346	0.000594
	0.10	0.000203	0.000402	0.000789
	0.30	0.000239	0.000410	0.000811
0.15	0.02	0.000469	0.000690	0.001092
	0.10	0.000308	0.000888	0.002662
	0.30	0.000502	0.001870	0.004763
0.50	0.02	0.000517	0.000654	0.000999
	0.10	0.000141	0.000361	0.001164
	0.30	0.000011	0.000132	0.002326
SR model				
<i>r</i>	<i>s</i>	$t = s/2$	$t = s$	$t = 2s$
0.02	0.08	0.000151	0.000320	0.000587
	0.30	0.000094	0.000159	0.000247
	0.60	0.000028	0.000060	0.000068
0.16	0.08	0.000266	0.000617	0.001725
	0.30	0.000243	0.000766	0.001525
	0.60	0.000144	0.000299	0.000478
0.50	0.08	0.000173	0.000388	0.001072
	0.30	0.000012	0.000102	0.001058
	0.60	0.000009	0.000139	0.000742

Initial state: repulsion gametes, $N = 50$.

model. Also as expected, an increase in fitness of the double mutant increased P in both models (Table 2).

In agreement with the first-generation analysis, asymmetries in fitness of the intermediate genotypes did not affect P in the RS case. For example, for $s =$

$t = 0.6$ and $r = 0.02$, P was 0.000472 in the absence of asymmetries and 0.000468 when $\alpha = 0.4$. In the SR model, asymmetries did have an effect for relatively large values of s . For example, for $s = t = 0.6$ and $r = 0.02$, P was 6×10^{-5} in the absence of asymmetries and 3.7×10^{-5} when $\alpha = 0.4$.

Finally, it is possible to express the probability of fixation as a function of the probability that at least one coupled gamete remains in the population after the first generation starting with the two single mutants (calculated using eqns. 4 and 5), and the probability of fixation of *ab* once the double mutant exists (given in Fig. 1). The results show a very good agreement for strong selection coefficients especially for the SR model (good agreement for $s = 0.7$ or larger in the SR model, for $s = 0.9$ and larger for the RS model; results not shown).

The time to fixation decreased monotonically with s and was independent of r in both models (Fig. 4). Asymmetries did not affect T (results not shown) while the fitness of *ab* had a very strong effect. For example in the RS model for $r = 0.02$ and $s = 0.1$, T was 79.5 generations, for $t = 0.05$ and only 40.8 generations for $t = 0.2$. The results on T support the idea that recombination and the fitness of the intermediate genotypes have a strong influence on the probability of generation of the double mutant. For fixation to occur, the double mutation gamete has to appear within the first few generations after the appearance of the single mutants. Once the double

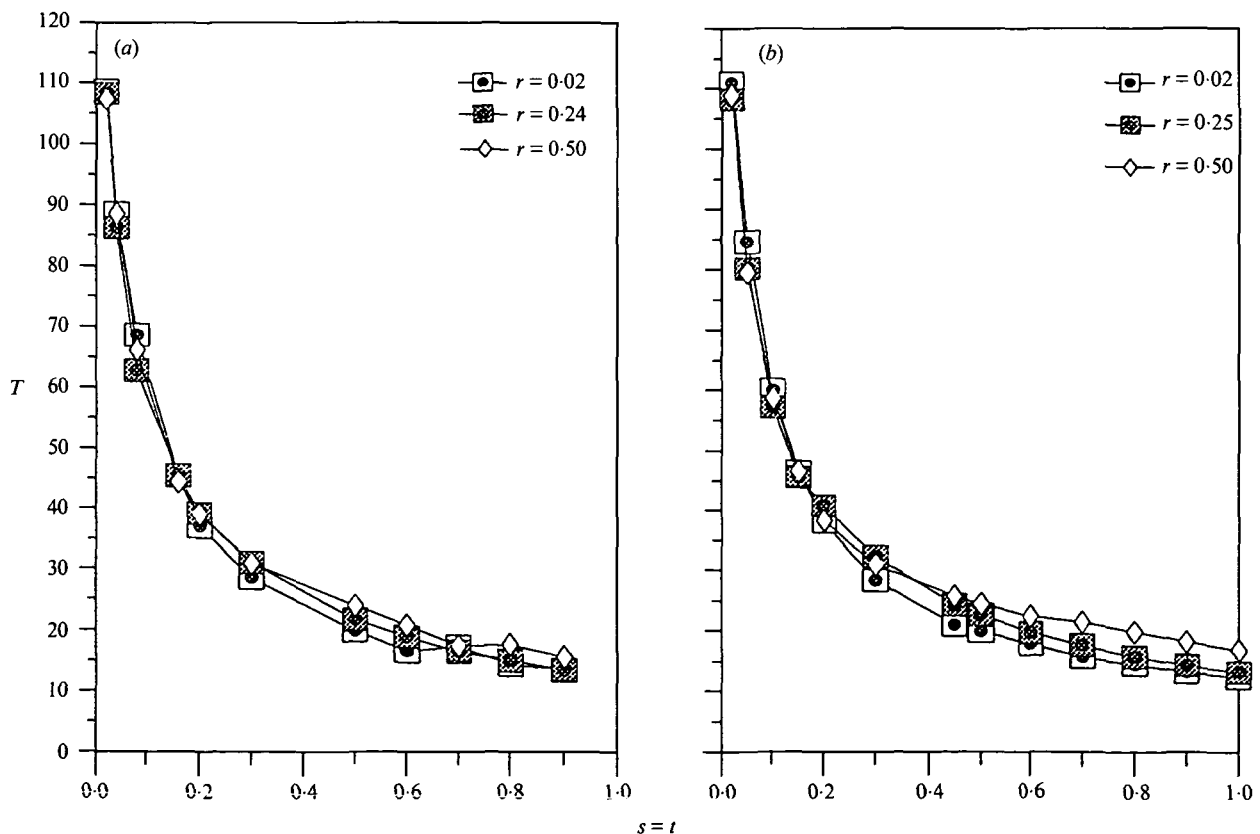


Fig. 4. Time to fixation of *ab* when the initial population contains one copy of each repulsion gamete. (a) SR model, (b) RS model.

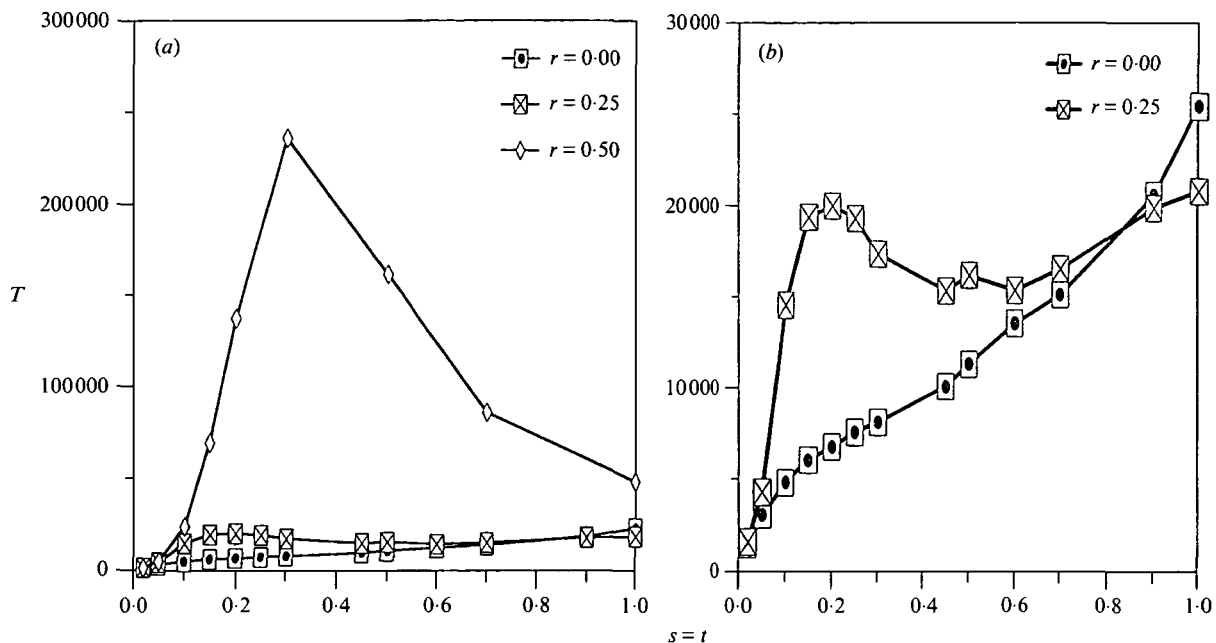


Fig. 5. Time to fixation of ab for RS model. The population is initially fixed for AB and variability is introduced by mutation at a rate $\mu = 0.001$ per generation. Panel (b) is a blow up of the results for $r = 0.00$ and $r = 0.25$. 1000 simulations per data point.

mutant is generated, the fate of the process mainly depends on the fitness of the double mutant.

(i) Effect of population size

The same qualitative results hold for the other values of N that we used (i.e. $N = 25$ and 100). Quantitatively, P always decreased as N increased, and the values obtained differed in order of magnitude for all values of r and strength of selection.

4.2. Dynamic model

Time to fixation as well as all the other statistics were estimated by running 1000 simulations for each parameter value combination for $N = 50$, and 200 simulations for each parameter value combination for $N = 25$. For $N = 100$ we ran 1000 simulations for each value of $s = t$ for $r = 0.0$ and 0.25 , and 200 simulations for $r = 0.5$. The only exception was for $N = 100$, $r = 0.5$ and $s = 0.25$ for the RS model where we ran only 100 simulations because of the excessively long time to fixation (approximately 3.8 million generations).

We estimated the time to fixation for three recombination rates, with a mutation rate $\mu = 0.001$. This mutation rate is large enough that the duration of the simulations is not prohibitive, but small enough that the mutation pressure is not dominating the process.

The SR and RS models behaved in very similar ways. We will illustrate the main features of the results using those obtained for the RS model, and will discuss the differences, if any, between the two models later.

Figure 5 shows the results we obtained for the RS

model. There are several interesting features. First, it should be noted that the times to fixation are generally very long given the unrealistically high mutation rate that we used. These times become extremely long if the mutation rate is further reduced. For example, the time to fixation with $\mu = 0.001$, $N = 50$, $r = 0.00$ and $s = t = 0.10$ and 0.25 was 4863.9 and 7668.9 generations respectively for the SR model. With the same parameter values except for a more realistic mutation rate ($\mu = 10^{-5}$) T was approximately 30 and 98 million generations respectively (average over 15 simulations for each case). It is worth noting, however, that when selection is weak, shifts happen reasonably often in agreement with the results of Barton & Rouhani (1987).

In agreement with Takahata's (1982) results, for a given selection coefficient the time to fixation is always larger as the recombination rate increases. This is especially true for intermediate selection coefficients.

The shape of the curve changed as the recombination rate increased. Time to fixation increased monotonically with s for low recombination rates, while the curve had a maximum at intermediate selection coefficients for large recombination rates. In other words, when the recombination rate was sufficiently large, mutations with strong effects, both detrimental when in repulsion and beneficial when in coupling, were fixed more rapidly than mutations with intermediate effects. This happened because at intermediate values of s , even though the repulsion gametes generate the double mutant relatively easily through recombination, selection in favour of the double mutant is not sufficient to drive it to fixation (see Figs. 1 and 3b). When s increased, the intermediate genotypes were removed from the population much

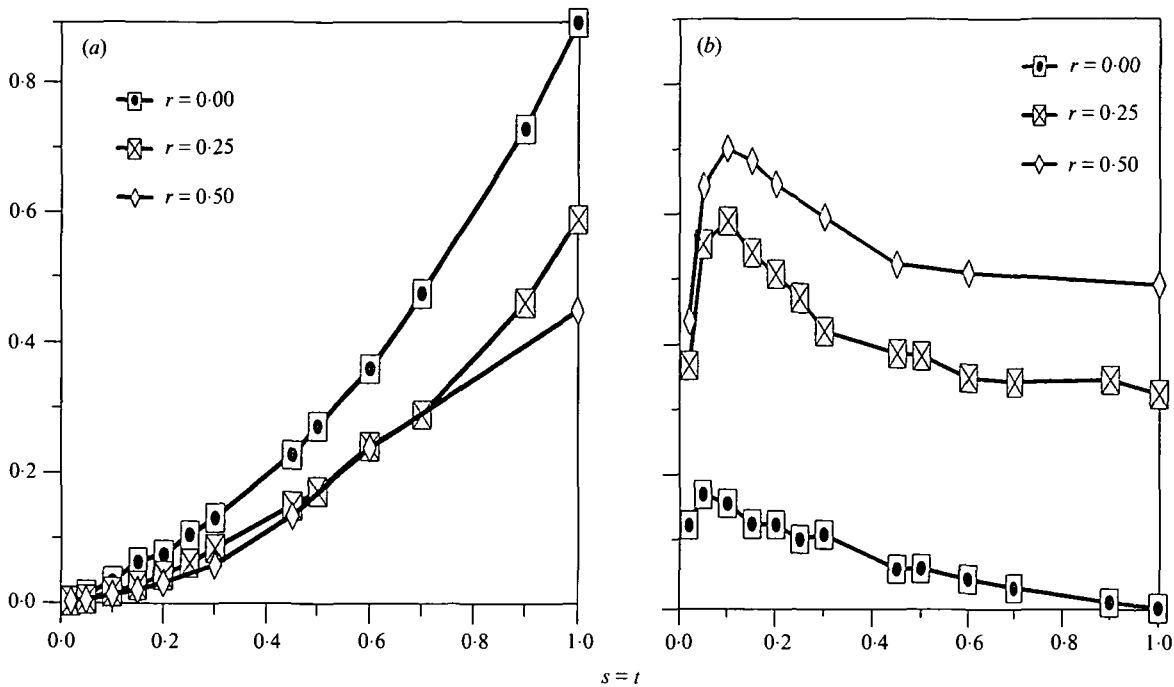


Fig. 6. Percentage of time that the double mutant appeared while the population was fixed for (a) *AB* (coupled gamete pathway) or (b) polymorphic at both loci (both-loci-polymorphic pathway). A locus is considered polymorphic if the rarest allele is present at at least one copy. Percentages are averaged over 1000 simulations. RS model.

faster, so that generation of *ab* via recombination was much more rare. The probability that the double mutant once generated is finally fixed, however, is much higher in this case.

In order to evaluate the importance of each pathway to the process of fixation of *ab*, we recorded the state of the population each time a double mutant appeared. We distinguished four situations: the population was fixed for *AB* and therefore the double mutant appeared directly as a coupled gamete (C); the population was fixed for one of the two intermediate genotypes (*Ab* or *aB*), the independent substitution path (IS); the population was polymorphic at both loci (P); and the population was fixed for any of the two alleles at one locus and polymorphic at the other (FP). The percentage of the first (C) and third (P) paths during the fixation process is shown in Fig. 6. As s increased the relative importance of the coupled-gametes path increased exponentially for all recombination rates. Furthermore, the lower the recombination rate the more important this pathway was, reaching 90% for $r = 0.0$ and $s = 1.0$. The pathway with both loci polymorphic (P) was a non-monotonic function of s , and was more frequent as r increased. Yet even when $r = 0.0$, when *ab* cannot be generated by recombination, this pathway is followed, because of the relatively large mutation rate used.

The independent-substitutions path lead to fixation only for small values of s ($s < 0.1$). This behaviour was predicted from the static case, because the probability of fixation of the intermediate genotypes was 0 for $s > 0.1$. For $s < 0.1$ the importance of the independent substitutions decreased rapidly. For example, for $r =$

0.0 the proportion of IS was 0.400 for $s = 0.02$, 0.084 for $s = 0.05$ and 0.001 for $s = 0.10$. For larger recombination rates these percentage values were slightly lowered (results not shown).

(i) *Effect of asymmetry in the fitnesses of intermediate genotypes*

Asymmetries in the fitnesses of intermediate genotypes always decreased the time to fixation. For example for $r = 0.5$ and $s = 0.05, 0.30$ and 1.0 the times to fixation were $3.2 \times 10^3, 1.4 \times 10^5$ and 3.0×10^4 generations respectively for $\alpha = 0.4$ as compared to $4.9 \times 10^3, 2.4 \times 10^5$ and 5.1×10^4 for $\alpha = 0.0$. This trend was true for $r = 0.25$ and $r = 0.00$ as well (results not shown). This result was not predicted by the static model. Indeed, under the static model asymmetries did not affect either the probability of fixation of *ab* or the time to fixation. This discrepancy may be explained by the difference in the way variability is introduced in the two models. In the static case, variability is part of the initial composition, and as soon as one of the two loci is fixed the population is 'trapped'. Asymmetries in this case do not affect the process because it is dominated by the fitness of the double mutant. In the dynamic model variability builds up in a more 'natural' way through mutation. Asymmetries in fitness of the two intermediate genotypes affect the rate of the process because, even though they do not affect the fate of the double mutant once it is formed, they do affect the rate at which it is formed. Indeed, in this case the intermediate genotype which has the highest fitness may be maintained in the population at

some intermediate frequency and generate the double mutant by mutation more easily than in the case with no asymmetries. This scenario is supported by the comparison of the relative importance of the different pathways. For example, for $r = 0.50$ and $s = 1.00$ the double mutant appeared through pathway C in 45.5%, pathway P in 49.9% and pathway FP in 4.6% of the cases, in the absence of asymmetries. With the same parameter values but $\alpha = 0.4$ the relative importance of the various pathways was 27.1%, 50.4% and 22.5% respectively. In other words, asymmetries change the way the population generates the double mutant, from the coupled-gametes pathway to the pathway where only one of two loci is polymorphic. In those cases where selection is weak, asymmetries increased the influence of the independent-substitutions pathway and decreased that of the other pathways (results not shown).

(ii) *Decoupling the positive and negative effects of the mutants*

The fitness of the double mutant had a very strong influence as expected. For a given selection intensity against the intermediate genotypes, the time to fixation decreased substantially as the fitness of the double mutant increased. For instance, for $r = 0.50$ and $s = 0.30$ the time to fixation was 2.4×10^5 for $t = 0.30$ while it was only 4.4×10^3 generations for $t = 1.00$. On the other hand, for a given value of the fitness of the double mutant (t) the time to fixation increased as selection against the intermediate genotypes increased. For example, for $r = 0.50$ and $t = 1.00$, T_{ab} was 2.7×10^2 , 4.0×10^2 , 4.4×10^3 , 2.4×10^4 and 5.1×10^2 generations for $s = 0.02$, 0.05 , 0.30 , 0.70 and 1.00 respectively. Phillips (1996) studied the effects of selection against the intermediate genotypes much more thoroughly.

The interaction between t and s is interesting for another reason. When selection against intermediate genotypes is weak and selection in favour of the double mutant is strong, Takahata's (1982) result, namely that the time to fixation increases with recombination rate, is no longer valid. In these cases the double mutant is fixed more rapidly in the presence of recombination than in the absence of recombination. For example, with $t = 1.00$ and $s = 0.02$ and 0.05 the average time to fixation was 271.2 and 400.7 generations for $r = 0.50$ while it was 386.7 and 657.4 generations respectively for $r = 0$. This happens because under weak selection against the intermediate genotypes, these genotypes are not immediately removed from the population. With $r = 0$, however, they cannot generate the double mutant by recombination. The population still has to wait till the double mutant is created by mutation. As the recombination rate increases another path is opened to the population, since the double mutant can now be

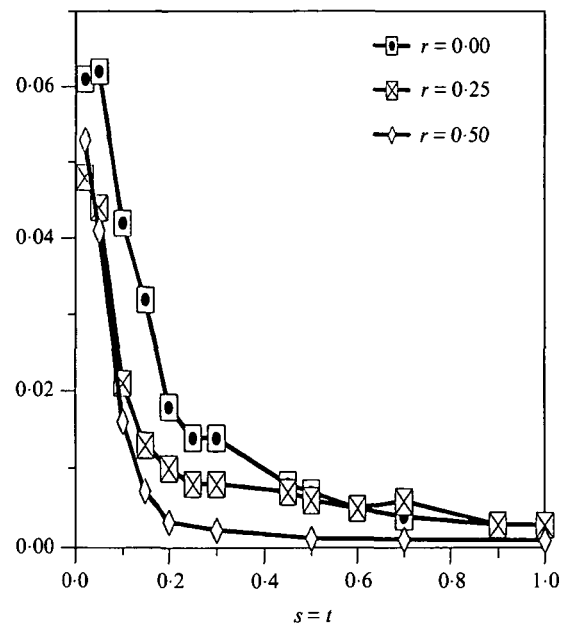


Fig. 7. Percentage of time that the population is polymorphic at both loci. A locus is considered polymorphic if the frequency of the rarest allele ≥ 0.05 . RS model.

generated by recombination. The process is thus shortened.

(iii) *Proportion of time the population is polymorphic at both loci*

The percentage of time the population was polymorphic at both loci is shown in Fig. 7. It is very low throughout the process, ranging from around 5% with very weak selection, decreasing very rapidly to around 1% for $s = 0.2$. Interestingly, for a given selection coefficient the population is polymorphic at both loci a smaller percentage of time as the recombination rate increases. This result indicates that one could conclude that variability in fitness is due to negative epistatic interactions only in a very small fraction of time. Actually this fraction would be further reduced by the fact that the frequency of both the intermediate genotypes and the double mutant are very small during the whole fixation process, so that the probability of actually sampling all kinds of alleles within a population, given that the population is polymorphic, is relatively small. Finally the percentage of time the population is polymorphic at both loci would be even smaller, in fact negligible, if the mutation rate was more realistic. For example, in 15 simulations we ran with the SR model, $N = 50$, $s = t = 0.10$ and $\mu = 10^{-5}$ the percentage of time the population was polymorphic at both loci was 6.6×10^{-6} .

For the sake of comparison we obtained equivalent results (with $\mu = 10^{-3}$) for an additive model. In this case the population is polymorphic at both loci

approximately 6% of the time. The time to fixation is also much faster under the additive model. We compare the percentage of time a population would be polymorphic at both loci under the two types of selection, additive and negative epistatic, for a given time interval by dividing each percentage by the corresponding average time to fixation. For $t = 0.1$ this ratio was 20.4 and 241.4 for $r = 0.0$ and 0.5 respectively, while for $t = 0.5$ the ratio was 756.1 and 97435.1 respectively. These results clearly indicate that most variability in fitness will be attributed to additive interactions, even though negative epistatic interactions might be present.

(iv) Effect of population size

Changes in population size affected the process in a predictable way. For a given value of Ns and Nr the process is enhanced as N increases, because $N\mu$ increases. For example for $Nr = 25$ and $Ns = 1$, T was 1.4×10^3 for $N\mu = 0.1$ and 1.5×10^3 for $N\mu = 0.05$, for $Nr = 25$ and $Ns = 10$, T was 2.8×10^4 for $N\mu = 0.1$ and 1.4×10^5 for $N\mu = 0.05$. The qualitative patterns were not affected by increases of N (results for $N = 25$ and 100 not shown).

The effect on T of increasing N while keeping r and s constant depended on the relative magnitudes of r and s . In the absence of recombination T decreased as N increased for all values of s (results not shown). With free recombination T increased as N increased for $s \leq 0.5$, while T decreased as N increased for $s = 1.0$. For an intermediate value of r ($r = 0.25$) the model behaved in an intermediate way: T decreased as N increased for $s \geq 0.5$, it increased as N increased for $s = 0.10$, while for $s = 0.25$, T was maximized for intermediate values of N . Hence, when selection is more important than recombination, increasing N decreases T . When recombination is more important than selection the outcome is the opposite: increasing N increases T .

Finally, as N increases the percentage of time the population is polymorphic at both loci increases (but remains below 10% even with $N = 100$), and the proportion of times that the double mutant first appears through the coupled-gamete pathway (C) increases (results not shown).

(v) Effect of the life-cycle (results not shown)

Not surprisingly, when one of the main two processes, recombination or selection, is weak the RS and SR models yield very similar numerical values. Thus, in the absence of recombination or when selection is weak ($s = t < 0.20$) the results were almost identical.

When both processes have an appreciable effect, fixation is faster under the SR model when recombination is stronger than selection (i.e. $r = 0.5$, $s = t \leq 0.5$). The opposite is true when selection is stronger than recombination (i.e. when $r = 0.25$ for all

values of $s = t$, and when $r = 0.50$ for $s = t \approx 1.0$). In the former case (recombination stronger than selection), recombination destroys the double mutant very rapidly. Fixation is faster when selection acts before recombination, because the frequency of the double mutant is increased by selection before recombination may break it down. In the latter case (selection stronger than recombination) the generation of the double mutant is much more limiting than its fate once it is present in the population. The RS model leads to a faster fixation because it allows the double mutant to be generated more frequently.

Finally, it should be noted that even when differences exist between the two models, the numerical results are of the same order of magnitude. The only exception is when recombination is stronger than selection and N is relatively large (i.e. $N = 100$) where the time to fixation of the RS model is approximately twice as large as that of the SR model. For example, when $N = 100$, $r = 0.5$ and $s = t = 0.25$, T was 3.8×10^6 generations for the RS model while it was only 1.6×10^6 generations for the SR model.

The life cycle does not affect the proportion of time the population is polymorphic at both loci, while not surprisingly the percentage of times that fixation follows the coupled-gamete pathway (C) increases if selection acts before recombination as soon as recombination is present (results not shown).

In summary, we have shown that the time to fixation increases as the fitness of the double mutant decreases and as selection against the intermediate genotypes increases. Fixation is faster in general when the recombination rate is small, except when selection against intermediate genotypes is weak and selection in favour of the double mutant is very strong. In these cases fixation is faster when the recombination rate is large. Mutations of strong effects are fixed more easily than mutations of intermediate effects, at least for large recombination rates. Independent substitutions lead to the fixation of the double mutant only when selection is weak. The relative importance of the other pathways depends on the interaction between recombination and selection, the coupled-gamete pathway being more influential as selection increases and recombination is reduced. Asymmetries in fitness between the intermediate genotypes increase the rate at which fixations occur, for all recombination rates. Finally throughout the fixation process, the population will be monomorphic at least at one of the two loci for most of the time, such that the opportunity to actually detect the presence of negative epistasis will be relatively scarce.

5. Discussion

We have considered how selection and recombination interact in the fixation of epistatic gene complexes. The static model examines the fixation process when one copy of each mutant is initially present. We

showed that the initial state of the population has a large effect on the probability of fixation of the double mutant. When the double mutant is initially present, and for that matter as soon as the double mutant is present whatever the initial state of the population, the probability of fixation is maximized by reduced recombination rates and large selection coefficients. When the mutant alleles are present under the form of repulsion gametes, the probability of fixation is maximized by intermediate recombination rates. Some recombination is needed to generate the double mutant, but not too much because recombination also tends to break up that genotype once it is generated. The strength of selection that maximizes P depends on the life-cycle and on the way selection against intermediate genotypes and selection in favour of the double mutant are linked, if any.

The static model is useful because it allows us to identify the factors that are important for the fixation process once the necessary genetic variability is present. The dynamic model is, however, more realistic because in it genetic variability is introduced by recurrent mutation. In the dynamic model, fixation of the double mutant is certain and the question is how long it takes. When selection against the intermediate genotypes (Ab and aB) is as strong as selection in favour of the double mutant (ab), the time to fixation depends in a complex way on the recombination rate and selection intensities. For small values of r , the time increases with the strength of selection. For larger recombination rates, there is an intermediate optimum in selection intensity. Stronger selection on the double mutant will decrease the fixation time.

In general, time to fixation increases as the recombination rate increases. This increase is not gradual, and it depends on the selection coefficient. For small selection coefficients intermediate recombination rates behave as large r since T is much larger in the presence of recombination than in the absence of recombination. For large s , however, intermediate recombination rates behave rather like the small recombination rates. In all cases, however, the process of fixation of the double mutant is much slower with free recombination than in the absence of recombination. This result was already found by Takahata (1982) who investigated the advantages or disadvantages of recombination for the accumulation of mutant genes.

The only case where recombination enhances the fixation of gene complexes is when selection against intermediate genotypes is very weak while selection in favour of the double mutant is very strong. This is because under weak selection against repulsion gametes, recombination may generate the double mutant which is then strongly selected for. In the absence of recombination, the only process which may generate the double mutant is mutation.

The pathway to fixation of the double mutant depends on the interaction between selection and

recombination. When selection is weak, independent substitutions account for approximately half of the fixations but the proportion decreases as the strength of selection increases. The population reaches fixation through the coupled-gametes pathway as selection intensity increases and the recombination rate decreases. In that case, the population mainly waits till a double mutant appears by mutation and then drives it to fixation. The importance of the recombination pathway, where both loci are polymorphic when the double mutant appears, decreases very slowly as the strength of selection increases but decreases sharply as the recombination rate is reduced.

The population is polymorphic at both loci only a small fraction of the time before fixation of the double mutant. Thus the chance of detecting epistatic interactions using methods requiring within populations polymorphism would be very small. It is much more likely that deleterious epistatic interactions may be revealed by methods requiring between populations (or even higher level) polymorphism. Indeed, the only studies evidencing deleterious epistatic interactions concerned interspecific crosses (see Palopoli and Wu, 1994 and references therein).

6. Conclusion

Several authors have argued that the third phase of Wright's shifting-balance theory, that is the spread of a mutant genotype once it is fixed in one population, will take place under a broad range of conditions (Crow *et al.* 1990; Phillips, 1993; Kondrashov, 1992; though see Barton, 1992). The feasibility of the process therefore depends on the initial phase, the fixation within one population.

Our results, in agreement with Phillips' (1996) results, indicate that even though the shift from one peak to another is possible, it takes a long time even with the unrealistically high mutation rates we used. Additive interactions make more rapid the fixation of favourable gene complexes. Therefore, if several alternative paths are available to evolution, the path of negative epistatic interactions probably does not occur frequently. There must exist some kinds of genetic changes, however, where there is no choice. In these cases the fixation of gene complexes will in general be faster if recombination is reduced. It is interesting to notice that under loose linkage, mutations of strong effects are fixed easier than mutations of intermediate effects.

We do not believe that negative epistatic interactions are the predominant mechanism in the fixation of favourable gene complexes. But important evolutionary transitions including ones that could not occur by other means could occur this way. Our results show that the absence of epistatic interactions found in an experimental analysis does not preclude the occasional success of the shifting balance process.

We thank Patrick Phillips for sharing unpublished results and discussions. Mike Whitlock provided helpful and stimulating discussions. Nick Barton and Sally Otto provided helpful criticisms. Drazen Borcic helped us with the 3-D graphs. This research was supported in part by CNRS to Y.M. and by a grant from the US National Institutes of Health (GM-40282) to M.S. Part of this research was done while M.S. was supported by the CNRS of France and visiting the laboratory of François Bonhomme in Montpellier, France, while another part was done while Y.M. was visiting the laboratory of Steve Frank in Irvine, California, and was supported by National Science Foundation grant BSR-9057331 and National Institutes of Health grants (GM-40403 and BSRG-S07-RR07008) to S. Frank.

References

- Barton, N. H. (1992). On the spread of new gene combinations in the third phase of Wright's shifting-balance. *Evolution* **46**, 551–557.
- Barton, N. H. & Rouhani, S. (1987). The frequency of shifts between alternative equilibria. *Journal of Theoretical Biology* **125**, 397–418.
- Barton, N. H. & Rouhani, S. (1993). Adaptation and the 'shifting balance'. *Genetical Research* **61**, 57–74.
- Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. New York: Harper and Row.
- Crow, J. F., Engels, W. R. & Denniston, C. (1990). Phase three of Wright's shifting-balance theory. *Evolution* **44**, 233–247.
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. London: Dover.
- Kimura, M. (1985). The role of compensatory neutral mutations in molecular evolution. *Journal of Genetics* **64**, 7–19.
- Kondrashov, A. S. (1992). The third phase of Wright's shifting-balance: a simple analysis of the extreme case. *Evolution* **46**, 1972–1975.
- Moore, F. B.-G. & Tonsor, S. J. (1994). A simulation of Wright's shifting-balance process: migration and the three phases. *Evolution* **48**, 69–80.
- Palopoli, M. F. & Wu, C.-I. (1994). Genetics of hybrid male sterility between *Drosophila* sibling species: a complex web of epistasis is revealed in interspecific studies. *Genetics* **138**, 329–341.
- Phillips, P. C. (1993). Peak shifts and polymorphism during phase three of Wright's shifting-balance process. *Evolution* **47**, 1733–1743.
- Phillips, P. C. (1996). Waiting for a compensatory mutation: phase zero of the shifting-balance process. *Genetical Research* **67**, 271–283.
- Rutledge, R. A. (1970). The survival of epistatic gene complexes in subdivided populations. Unpublished PhD Thesis, Columbia University.
- Takahata, N. (1982). Sexual recombination under the joint effects of mutation, selection and random sampling drift. *Theoretical Population Biology* **22**, 258–277.
- Williams, G. C. (1992). *Natural Selection. Domains, Levels and Challenges*. New York: Oxford University Press.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.