Theoretical Population Biology **51**, 134–147 (1997)
Article No. TP971301

Deleterious Mutations, Variable Epistatic Interactions, and the Evolution of Recombination

Sarah Perin Otto*

Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4

and

Marcus W. Feldman

Department of Biological Sciences, Stanford University, Stanford, California 94305

Received February 20, 1996

In this paper, we examine the conditions that allow increased recombination to evolve in the presence of recurrent deleterious mutation. We focus on a three-locus model first studied by Feldman *et al*. (1980), which follows the dynamics of a modifier locus that alters the recombination rate between two loci subject to deleterious mutation. Although Feldman *et al*. (1980) indicated that increased recombination might be favored if there is diminishing-returns epistasis, we show that alleles that increase the recombination rate can only invade if there is synergistic epistasis between the loci under selection. Even with synergistic epistasis, evolution at the modifier locus will lead to *decreased* recombination if the modifier locus is loosely linked and epistasis is strong. Using the multi-locus analysis of Barton (1995), we show that variability among loci in the sign and strength of epistasis further decreases the parameter space over which increased recombination may evolve. We conclude that, even with negative epistasis, increased recombination may only be favored when linkage is tight, especially if, as seems likely, epistatic interactions are highly variable among loci.

1. INTRODUCTION

By reducing genetic associations among loci, recombination can exert a strong influence on the genetic changes that occur within a population in response to selection (reviewed in Michod and Levin, 1988, and Kondrashov, 1993). Recombination has traditionally been thought to have a beneficial influence on the efficiency of selection in a variety of different contexts: recombination can generate new beneficial genetic combinations (Fisher, 1930; Muller, 1932); it can assist in the tracking of fluctuating environments (Sturtevant & Mather, 1938; Charlesworth, 1976; Hamilton, 1980); and it can hasten the elimination of deleterious mutations (Muller, 1964; Kimura & Maruyama, 1966; Feldman *et al*., 1980; Kondrashov, 1982). In each of these cases, increased recombination may, under the appropriate conditions, have a selective advantage (Barton, 1995; Feldman *et al*., 1997; Charlesworth, 1990; Charlesworth, 1993). As a general explanation for the evolution and maintenance of sex and recombination, the hypothesis

* To whom correspondence should be addressed. E-mail: otto@zoology.ubc.ca.

0040-5809/97 $25.00

that recombination aids in the elimination of deleterious mutations from the genome is especially compelling, since deleterious mutations pose such a common and persistent problem to all living organisms (Kondrashov, 1988; Kondrashov, 1993). In this paper, we will examine the evolution of recombination in a large population subject to recurrent deleterious mutations.

Mutations arise continuously despite the existence of several mechanisms that limit DNA damage and repair replication errors. Unfortunately, data on the rate of deleterious mutations across a genome are limited. Minimum estimates of the average number of new deleterious mutations per diploid genome per generation, U_{min}, are approximately $\frac{1}{2}$ in *Drosophila melanogaster* (Houle *et al*., 1994; Keightley, 1994; Mukai *et al*., 1972), $\frac{1}{2}$ in a number of inbred flowering plant species (Charlesworth *et al*., 1990), one in humans (Kondrashov & Crow, 1993), and 0.0002 in *Escherichia coli* (Kibota & Lynch, 1996). Purifying selection eliminates these mutations, but its efficiency depends upon several factors, including the epistatic interactions and recombination rates among mutant loci. In response to this process, the genetic system (e.g. linkage relationships) may itself evolve. One approach to modeling the evolution of the genetic system is to follow the dynamics of modifier alleles that alter some aspect of the genetic system such as the recombination rate or the mutation rate (Nei, 1967; Feldman, 1972; Feldman *et al*., 1980; Brooks, 1988).

In this work, we will attempt to synthesize known results from three-locus and multi-locus models that examine the evolution of recombination at a modifier locus in response to recurrent deleterious mutations. This attempt has made us aware of an error in a previous paper that examined the evolution of recombination in a three-locus model (Feldman *et al*., 1980). This error is corrected in the appendix to this paper. We then use both the three-locus and multi-locus models to address the importance that variable epistatic interactions may have on the evolution of recombination. Previous multi-locus analyses of recombination in the presence of deleterious mutations tend to assume specific functional forms (e.g. truncational or quadratic fitness functions) for the way loci interact to affect fitness (Kondrashov, 1984; Charlesworth, 1990). In particular, these models assume that all loci interact in an equivalent manner with the strength of epistatic interactions depending only on the number of previous mutations and not on their nature. Although little is known about the distribution of epistatic interactions among loci, it is quite likely that much variation exists in the form and strength of these interactions (Whitlock *et al*., 1995), with some loci interacting strongly and others weakly. We find that variation in epistasis tends to favor reduced recombination. Variable epistatic interactions thus limit the parameter space in which increased recombination evolves in response to recurrent deleterious mutation.

The importance of epistasis. If the fitness of an organism, W, were simply the product of the fitness, $(1-s_i)$, of each mutation carried within the genome, i.e. if fitnesses were multiplicative and $W=\prod_{i=1}^{n}(1-s_i)$, then selection would not create genetic associations among loci. Between two loci, for example, linkage disequilibrium is absent at an equilibrium between mutation and multiplicative selection (Feldman *et al*., 1980). When genetic associations are absent, the level of recombination becomes immaterial since haplotypes cannot be further randomized (Maynard–Smith, 1968). This result breaks down, however, with epistatic interactions among loci (Eshel & Feldman, 1970). Epistasis (assumed in this paper to measure the departure from multiplicative selection) ensures that linkage disequilibrium will be maintained at a mutation-selection balance (Feldman *et al*., 1980). These genetic associations allow indirect selection to act on modifier alleles that alter recombination rates.

Unfortunately, little empirical data exist about the epistatic relationships among random mutations. The data that best serve as an empirical guide come from experiments examining the effects of mutations on viability in *D. melanogaster* (Mukai, 1964; Mukai, 1969; Mukai *et al*., 1972). In these experiments, selection against deleterious mutations on the second chromosome was relaxed, allowing mutations to accumulate over time. From these experiments, Mukai *et al.* (1972) estimated that the rate of deleterious mutation on the second chromosome is, at a minimum, 0.12 per generation, leading to an estimate for the rate of deleterious mutations across the diploid genome of $U=0.6$. The average selection coefficient against mutations was, at a maximum, $\bar{s}\approx 0.038$, with a mean dominance coefficient of $\bar{h}\approx 0.21$. In the Mukai *et al.* (1972) study, mean viability went down linearly over 40 generations, with no significant additive or multiplicative epistasis. A mutation accumulation experiment in *E. coli* also failed to find evidence of epistatic interactions among accumulated mutations (Kibota & Lynch, 1996). In an earlier study that extended through 60 generations, however, Mukai (1969) observed a significant quadratic relationship among mutations in *D. melanogaster*, such that the deleterious effects of mutations were even greater in combination ("synergistic" or negative epistasis). Crow (1970) fitted a quadratic fitness function to these data, finding that the fitness of an

individual bearing *n homozygous* mutations fitted the relation

$$W(n) = 1 - an - bn^2 = 1 - 0.014n - 0.011n^2. \quad (1)$$

Here the quadratic term measures the extent of additive epistasis (departure from linear selection). To measure the extent of multiplicative epistasis among mutations that appear as *heterozygotes*, Charlesworth (1990) used an estimate of dominance of $h = 0.2$, Crow's formula, and Mukai's data to find that approximately

$$\ln[\,W(n)] = -\alpha n - \tfrac{1}{2}\beta n^2 = -0.002n - 0.0004n^2, \quad (2)$$

where β measures the departure from multiplicative selection. The general applicability of these fitness functions remains to be demonstrated; in particular, confirmation is needed that deleterious mutations do tend to interact synergistically and that their interactions can be described by a quadratic fitness function.

A point to remember is that the quadratic fitness function was estimated using the average viability of seventy-two lines of *D. melanogaster*. Each of these lines accumulated a different sequence of mutations, and it is likely that the actual fitness functions for the different lineages were quite different, with some lines exhibiting stronger and others weaker epistatic interactions. Certainly, substantial variation in viability was observed among the lines (the mean viability after 60 generations was 0.51 with a standard deviation of 0.25), but this variability could have had several causes including (1) variation in the number of mutations accumulated (2) variance in the selective strength of these mutations, and (3) variance in the epistatic interactions among them. No data currently exist on the variance in epistatic interactions among random mutations. Nevertheless, the wide variety of ways in which gene products interact at a molecular level strongly suggests that epistatic interactions do vary dramatically depending on the loci involved (Whitlock *et al.*, 1995).

MODELS FOR THE EVOLUTION OF RECOMBINATION WITH EPISTASIS

Three-locus model. A three-locus model was developed by Feldman *et al.* (1980) to investigate the evolution of recombination between two loci subject to recurrent deleterious mutations. We will focus on the simpler haploid model in the text, with an analysis of both the haploid and diploid models given in the appendix. In this model, one locus M modifies the recombination rate between two viability loci (A and B) as shown in Table 1. The gene order is assumed to be MAB, with the M locus located a distance, R, from the A locus. It is assumed that there is no interference among crossovers within this region. The model is unchanged if we assume that the modifier also changes the rate of recombination between itself and locus A. R, in this case, measures the rate of recombination between M and A in heterozygotes (Mm) at the modifier locus (the rate of recombination between M and A in MM and mm homozygotes has no effect on the dynamics of the system). Both viability loci are subject to recurrent deleterious mutations (to a and b) at a rate μ. Selection acts against these mutations, such that the viabilities of AB, Ab, aB, and ab are 1, $W_1 = 1 - s$, $W_1 = 1 - s$, and $W_2 = (1 - s)^2 + \varepsilon$, respectively. By assumption, additional mutations always decrease fitness ($1 > W_1 > W_2$), and both μ and s are the same for the A and B loci. When ε is negative, mutations have a more deleterious effect on fitness in combination than expected if fitnesses were multiplicative ("synergistic" or "negative" epistasis), whereas when ε is positive, mutations have a weaker effect on fitness when combined ("diminishing-returns" or "positive" epistasis). During reproduction, haploids fuse to form diploids, which then undergo recombination and mutation to regenerate the haploid stage. Censusing occurs at the adult haploid stage after selection.

With one modifier allele present, M say, the system approaches a two-locus mutation-selection balance as long as the mutation rate is small relative to the strength of selection ($(1 - \mu) > W_1$ and $(1 - \mu)\, W_1 > W_2$) (Feldman *et al.*, 1980). At this equilibrium, the linkage disequilibrium, $\hat{D}$, between the two viability loci has the same sign as ε (Feldman *et al.*, 1980). The stability of this equilibrium to invasion by a second modifier allele, m, was analysed by Feldman *et al.* 1980. One part of their analysis contains an error which is corrected in the appendix to this paper. Also in the appendix, we derive the stability conditions for weak epistasis and for free recombination between the modifier and the loci under direct selection. Here, we summarize the main findings.

TABLE 1

The Effect of a Modifier of Recombination

Modifier genotype	Recombination rate between A and B
MM	r_1
Mm	r_2
mm	r_3

Either with tightly linked modifiers (R small) or with extremely weak epistasis ($\varepsilon \ll \mu, s$), the sign of the epistasis governs evolution at the modifier locus. When epistasis is positive, modifiers that decrease the recombination rate are favored. When epistasis is negative, modifiers that increase the recombination rate are favored. Feldman *et al.* (1980) noted, however, that if linkage is sufficiently loose between the modifier locus and the loci under selection (such that R is above a critical value R^*), a switch in stability can occur. Although their analysis implied that such a switch could occur with both positive and negative epistasis, the reanalysis presented in the appendix indicates that there is no switch in stability if epistasis is positive for biologically relevant values of the recombination rate (within $[0, \frac{1}{2}]$); that is, decreased recombination is always favored with positive epistasis. When there is moderately strong *negative* epistasis, however, a critical value ($R^*; 0 < R^* < \frac{1}{2}$) may be reached such that for all $R > R^*$ increased recombination is no longer favored and modifiers that *decrease* recombination can invade. Consequently, for unlinked modifiers ($R = \frac{1}{2}$), increased recombination is only favored if there is fairly weak negative epistasis. From Eq. (19) in the appendix, which assumes that the mutation rates are much smaller than the selection coefficients, increased recombination evolves when $R = \frac{1}{2}$ only if

$$\frac{-s^2(3-s)}{(1-s)} < \varepsilon < 0. \tag{3}$$

As noted above, the results apply equally well if we assume that R is the modified recombination rate between M and A in Mm individuals. Thus, condition (3) must be met for the invasion of a modifier allele that, when heterozygous, produces free recombination between all loci, even if the modifier is tightly linked to the A locus in MM homozygotes. These results are consistent with an analysis of the three-locus special case of a more general multilocus model in Barton (1995). Barton's method assumes that a modifier has only a slight effect on recombination rates ($r_1 \approx r_2 \approx r_3$), an assumption which does not appear to be critical to the qualitative conclusions concerning the three-locus model.

A heuristic explanation for the behavior of the three-locus model. We have observed that recombination is never favored under positive (diminishing-returns) epistasis and is favored under negative (synergistic) epistasis when the modifier is tightly linked to the loci under selection or, if it is loosely linked, when epistasis is sufficiently weak. These phenomena can be explained in terms of the genetic associations that develop between a modifier and the selected loci when a modifier allele appears in a population (Barton, 1995; Bergman *et al.*, 1995). Assuming that the modifier allele is initially in linkage equilibrium with the selected loci, it can be shown (Barton, 1995, p. 130; see Bergman *et al.* 1995 for a similar proof) that genetic associations develop initially such that the modifier allele becomes positively associated with both the AB and ab genotypes and with both the A and B alleles if the modifier increases recombination and the epistasis is negative (or if the modifier decreases recombination and the epistasis is positive). This is expected on intuitive grounds. When epistasis is negative, AB is under-represented in the population ($\hat{D} < 0$) and modifier alleles that increase the recombination rate will have a higher proportion of offspring that are AB and ab. Among these offspring, selection will increase the frequency of the A and B alleles more effectively than in the remainder of the population (because of the greater variance in fitness) and so a modifier allele that increases recombination becomes positively associated with the A and B alleles.

Of course, the only alleles that will leave descendants in the long-term are the A and B alleles, so a genetic association with these alleles would cause a modifier to increase when rare. That is, genetic associations with the fittest alleles in the population would favor increased recombination in the presence of negative epistasis and decreased recombination in the presence of positive epistasis. As noted above, however, the modifier allele is not only associated with the fittest alleles but also with both the AB and ab genotypes (at least initially). The average fitness of these offspring will be greater than the average fitness of the other possible offspring, Ab and aB, only if $(1 + (1-s)^2 + \varepsilon)/2 > 1 - s$, which rearranges to $s^2 + \varepsilon > 0$. This condition is always met when epistasis is positive, so that a modifier that decreases recombination is favored both by genetic associations with the fittest alleles and by an increased fitness among its offspring. The condition is only satisfied, however, with negative epistasis if the epistasis is fairly weak $-s^2 < \varepsilon < 0$. With stronger negative epistasis ($\varepsilon < -s^2$), modifier alleles that increase recombination suffer from an immediate reduction in the average fitness of their offspring. In this case, producing the best and the worst genotype by recombination leads to a lower fitness on average than having the intermediate genotypes. When the modifier is tightly linked to the major loci, its genetic associations with the A and B alleles are strong enough to overwhelm the decrease in mean fitness caused by eliminating linkage disequilibrium. Loosely linked modifiers have,

however, only a weak genetic association with the fittest alleles, an advantage that is overwhelmed by the immediate fitness loss when, approximately, condition (3) is violated. In summary, decreased recombination is always favored when epistasis is positive. Increased recombination is only favored when epistasis is negative and when long-term associations with the fittest alleles outweigh immediate fitness losses due to the elimination of disequilibrium; this is made possible by weak epistasis or by tight linkage of the modifier to the selected loci. Interestingly, this verbal argument also explains the behavior of modifiers of recombination in models with directional selection (Barton, 1995; Bergman *et al.*, 1995).

A numerical example. As a concrete example, we can use Mukai's data (via Eq. (1)) to estimate the fitness of individuals carrying one mutation ($W_1 = 0.975$) and two mutations ($W_2 = 0.928$) so that $s = 1 - W_1 = 0.025$ and $\varepsilon = W_2 - W_1^2 = -0.023$. Although Eq. (1) was developed using data on the fitness effects of homozygous mutants, we will assume that the fitness effects of haploid mutants are equivalent. Since ε does not obey condition (3), we expect decreased recombination to be favored when the modifier is loosely linked. In fact, for R greater than approximately 0.05, only modifiers that decrease recombination are able to invade (Fig. 1).

In Mukai's experiment, the average fitness loss caused by a mutation becomes more severe after further generations of mutation accumulation. If we examine a population of individuals that already carry five mutations, the baseline fitness of the population is $W_5 = 0.655$ from (1). A sixth mutation has a much lower fitness, $W_6 = 0.520$, because of negative interactions with the previous five mutations. A seventh mutation experiences nearly the same fitness decline, $W_7 = 0.363$, since there is comparatively little difference between interacting with six rather than five loci. Let us measure the "apparent" strength of selection within this region as $s = 1 - W_6/W_5 = 0.206$ and the "apparent" epistasis as $\varepsilon = W_7/W_5 - (W_6/W_5)^2 = -0.076$. In this case, epistasis is sufficiently weak relative to the strength of selection for ε to obey condition (3) and so increased recombination is favored for all values of R (Fig. 2).

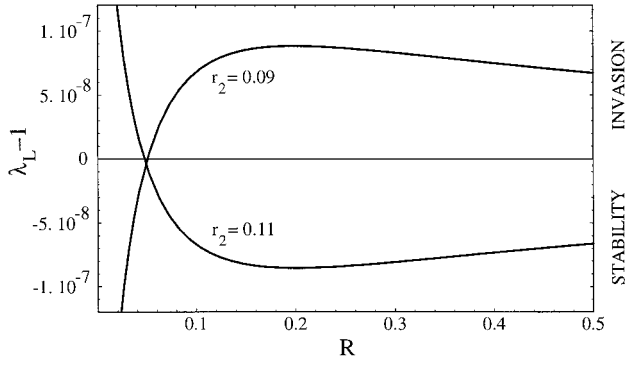

FIG. 1. Invasion of a modifier of recombination with $W_1 = 0.975$, $W_2 = 0.928$, and $r_1 = 0.1$. Invasion occurs when $\lambda_L - 1$ is positive. Invasion of a modifier that increases recombination (to $r_2 = 0.11$) is only possible when $R < 0.047$. Invasion of a modifier that decreases recombination (to $r_2 = 0.09$) occurs when $R > 0.050$. Even though neither of these modifier alleles can invade for $0.047 < R < 0.05$, weaker modifiers can invade in this range.

Correspondence with polygenic models. Many of the features observed in the three-locus model extend to the evolution of recombination in polygenic models (Kondrashov, 1984; Charlesworth, 1990). These models focus on the effects of recombination on the distribution of the number of mutations, n, carried by an individual. With synergistic epistasis, increased recombination is generally favored in the polygenic models, although decreased recombination evolves if (a) linkage is already loose, (b) epistasis is sufficiently strong, and (c) mutation rates are low. Of these three caveats, the first two were also observed with the three-locus model. The dependence on the mutation rate, however, is a little more subtle.

In the three-locus model, the stability conditions are fairly insensitive to the mutation rate. For example, Eq. (17) in the appendix depends only weakly on μ. How is it, then, that the magnitude of the mutation rate comes to play such a critical role in the polygenic models? The main effect of a high genomic mutation rate is to increase the average number of mutations carried by an individual at different loci (polygenic models assume a large number of contributing loci so that mutations at a locus are still rare even with high mutation rates). This

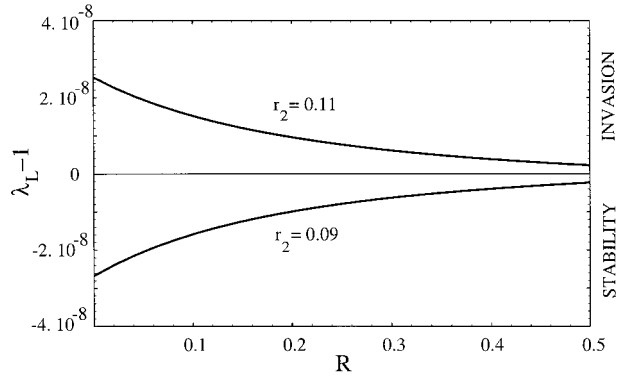

FIG. 2. Invasion of a modifier of recombination with $W_1 = 0.794$, $W_2 = 0.554$, and $r_1 = 0.1$. Invasion occurs when $\lambda_L - 1$ is positive. A modifier that increases recombination (to $r_2 = 0.11$) invades for all R.

changes the selective context in which a new mutation finds itself (Kondrashov, 1984). Generalizing from the previous section, if there are $\bar{n}$ mutations present in the genome on average, we can define the apparent strength of selection and epistasis at this point as

$$s = 1 - W_{\bar{n}+1/2}/W_{\bar{n}-1/2}$$
$$\varepsilon = W_{\bar{n}+1}/W_{\bar{n}-1} - (W_{\bar{n}}/W_{\bar{n}-1})^2. \quad (4)$$

Condition (3) suggests that a natural measure of the strength of epistasis is $\varepsilon^* = \varepsilon/s^2$, whereby condition (3) can be rewritten as $-3 \lesssim \varepsilon^* < 0$ for increased recombination to be favored if the modifier is unlinked. Using either fitness function (1) or (2) with negative epistasis (such that a, b, α, β are all positive), ε^* takes on its most negative value at $n = 0$ and rises rapidly and monotonically towards zero as n becomes large. For the fitness function (1) which describes the accumulation of homozygous mutants, there must be at least four mutations in the genome on average ($\bar{n} \geqslant 4$) for $\varepsilon^* > -3$ and hence for increased recombination to be favored at unlinked modifier loci. For the fitness function (2) which describes the accumulation of heterozygous mutants, $\bar{n}$ must be at least 18 for evolution to favor increased recombination at unlinked modifier loci. Since $\bar{n}$ increases with mutation rate, high genomic mutation rates will be associated with relatively weak negative epistasis (ε^* is small in magnitude), a condition that favors the evolution of recombination when linkage is loose.

Charlesworth (1990) and Charlesworth and Barton (1996) present a similar argument to explain why there is more often selection for increased recombination when mutation rates are high. They note that fitness functions such as (2) have an inflection point. When there are few mutations in the population (to the left of the inflection point), the fitness function has a negative second derivative, so increasing the variance in n (by recombination, for example, when there is synergistic epistasis) decreases mean fitness. Conversely, with many mutations in the population (to the right of the inflection point), there is a positive second derivative and increasing the variance in n increases mean fitness. Hence, modifiers that increase recombination are more likely to produce offspring with a higher mean fitness than the rest of the population if there are already many mutations within the population. These arguments are related; we simply point out that the relative strength of synergistic epistasis decreases as $\bar{n}$ increases and that this fact can be used to explain why recombination is more often favored with high mutation rates.

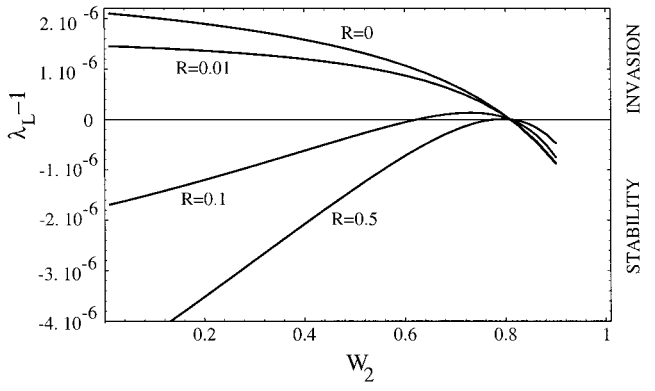

FIG. 3. Invasion of a modifier that increases recombination (from $r_1 = 0.2$ to $r_2 = 0.3$) with $W_1 = 0.9$. Invasion occurs when $\lambda_L - 1$ is positive. For $W_2 > 0.81$, there is positive epistasis and invasion does not occur. For $W_2 < 0.81$, there is negative epistasis and invasion does occur, but only when R is small and/or epistasis is weak (W_2 near 0.81).

Variance in epistasis. Condition (3) indicates that, when $R = \frac{1}{2}$, increased recombination is favored only for a small range of epistatic values. This point is further illustrated in Figures 3 and 4. Shown is the leading eigenvalue minus one, $\lambda_L - 1$, (obtained from a numerical evaluation of the stability matrix) as a function of the fitness of double mutants, W_2 (note the different axes). Invasion of the new modifier allele (here an allele that increases recombination between A and B from 0.2 to 0.3) occurs only when $\lambda_L - 1$ is positive. As can be seen from these graphs, increased recombination is favored over a very small parameter space when $R = \frac{1}{2}$. When $W_1 = 1 - s = 0.9$ (Fig. 3), the fitness of double mutants must lie between 0.78 and 0.81 for increased recombination to be favored. The situation is even less favorable for

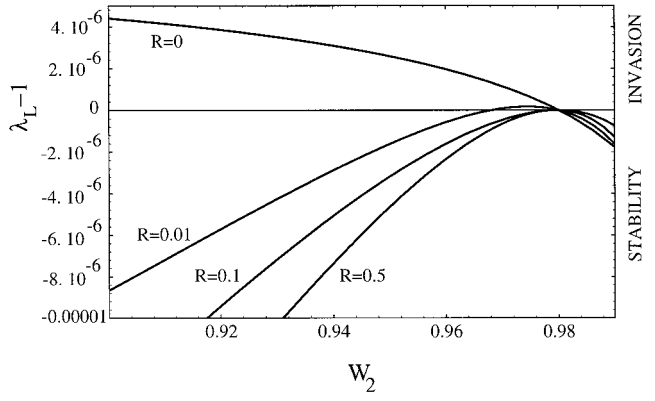

FIG. 4. Invasion of a modifier that increases recombination (from $r_1 = 0.2$ to $r_2 = 0.3$) with $W_1 = 0.99$. Invasion occurs when $\lambda_L - 1$ is positive. For $W_2 > 0.9801$, there is positive epistasis and invasion does not occur. For $W_2 < 0.9801$, there is negative epistasis and invasion does occur, but only when R is small and/or epistasis is weak (W_2 near 0.9801).

the evolution of recombination when selection is weaker. When $W_1 = 1 - s = 0.99$ (Fig. 4), the fitness of double mutants must lie between 0.9798 and 0.9801 for increased recombination to be favored. There must therefore be fairly small departures from multiplicative fitnesses for recombination to evolve when $R = \frac{1}{2}$.

The lower the recombination rate between the modifier and locus A the larger the range of W_2 over which increased recombination will be favored. Even when $R = 0.1$, increased recombination is favored only if W_2 lies between 0.65 and 0.81 when $W_1 = 0.9$ (Fig. 3) or between 0.9789 and 0.9801 when $W_1 = 0.99$ (Fig. 4). Similar graphs are observed for other modifier alleles (tested values include $r_1 = 0$ with $r_2 = 0.1$, $r_1 = 0$ with $r_2 = 0.01$, $r_1 = 0.2$ with $r_2 = 0.4$, $r_1 = 0.2$ with $r_2 = 0$), although the curves for $R = 0.01$ and $R = 0.1$ vary. The graphical analyses suggest that, as an approximate rule of thumb, R must be less than s for there to be a large range of values of W_2 in which evolution favors increased recombination.

When this range is small, the evolution of recombination may be quite sensitive to variation in the strength of synergistic epistasis. Even if the average pair-wise epistasis favors recombination, this average may be taken over a distribution of epistatic interactions which includes sets of loci that interact with strong negative epistasis (such that reduced recombination is favored unless the modifier is tightly linked) or with positive epistasis (such that reduced recombination is again favored). This suggests that variability in the strength of interactions among loci will tend to reduce the advantage of recombination. We can use the recent analysis of Barton (1995) to estimate the role of variable epistatic interactions on the evolution of recombination in a multi-locus model. Barton's (1995) Eq. (18a) gives the selection coefficient (s_i) acting upon a modifier of recombination at locus i, assuming that the effects of the modifier on recombination are weak (all terms used in this equation are defined in Table 2):

$$s_i \approx -\tfrac{1}{2} n^2 \, \mathrm{E}[\delta r_{jk} \varepsilon_{jk} \mu_j \mu_k] \, \mathrm{E}\left[\frac{1}{r_{ijk} r_{jk}} \left(\frac{1}{r_{ik}} + \frac{1}{r_{ij}} - 1\right)\right] - \tfrac{1}{2} \sum_{|S|>1} \delta r_{|S|} \, V_{|S|} \, \mathrm{E}\left[\frac{1}{r_{iS} r_S}\right]. \quad (5)$$

The first half of Eq. (5) quantifies the amount of indirect selection on a modifier that arises from changing the efficiency of selection (Barton, 1995). Recombination can increase the efficiency of selection by creating genotypes with very many and very few mutations if these genotypes are less common than expected (which they will be if there is negative epistasis). The second half of Eq. (5) quantifies the immediate fitness consequences to a modifier that arise from changed genetic associations among sets of loci (Barton, 1995). The second half of (5), as defined, always selects for decreased recombination; those genetic combinations that are more common than expected because selection against them is weaker than expected are destroyed by recombination. Assuming weak selection, weak epistasis, and a weak modifier, this formula is accurate as long as the expected recombination rate between loci is not very small, although Barton discusses reasonable approximations for this case.

TABLE II

Terms Used in Equation (5)

Term	Definition
n	Number of viability loci
E	The expectation over all distinct loci of the parenthetical quantities
ε_{jk}	Extent of multiplicative epistasis between loci j and k
μ_j	Mutation rate at locus j
S	A particular set of loci containing $\|S\|$ loci
$\|S\|$	The number of loci in a set S
$\sum_{\|S\|>1}$	The summation over all set-sizes with more than one locus
r_S	The probability that recombination breaks up a set of loci
r_{jk}	The probability that recombination occurs between loci j and k
r_{ijk}	The probability that recombination breaks up the set of loci i, j and k
r_{iS}	The probability that recombination breaks up the set of loci i and S
$\delta r_{\|S\|}$	The amount by which the modifier changes r_S
δr_{jk}	The amount by which the modifier changes r_{jk}
$V_{\|S\|}$	The variance in log fitness due to epistatic interactions among $\|S\|$ loci,
V_2	…measured as a departure from multiplicative epistasis
V_j	…measured as a departure from additive epistasis for $j > 2$

Any variability among sets of loci in the amount of pair-wise epistasis contributes only to the second half of Eq. (5), which, as noted by Barton, always favors decreased recombination. To simplify matters, assume that the mutation rate and the amount by which recombination is modified (δr_{jk}) are independent of the strength of epistasis at the relevant loci. Denote the average mutation rate by $\bar{\mu}$ and the average δr_{jk} between pairs of loci by $\overline{\delta r}$. Equation (5) then becomes

$$s_i \approx -\frac{1}{2} \overline{\delta r} (n\bar{\mu})^2 \, \mathrm{E}[\varepsilon_{jk}] \, \mathrm{E}\left[\frac{1}{r_{ijk} r_{jk}} \left(\frac{1}{r_{ik}} + \frac{1}{r_{ij}} - 1\right)\right] - \frac{1}{2} \overline{\delta r} \, V_2 \, \mathrm{E}\left[\frac{1}{r_{ijk} r_{jk}}\right] - \frac{1}{2} \sum_{|S|>2} \delta r_{|S|} \, V_{|S|} \, \mathrm{E}\left[\frac{1}{r_{iS} r_S}\right]. \quad (6)$$

In this equation, we have separated out terms that measure pair-wise interactions among loci (first two terms) from terms that measure interactions among three or more loci beyond their pair-wise interactions (last term). Although the last term always selects for decreased recombination, it is expected to be small (Barton, 1995), and we will ignore it in the following discussion.

V_2, which measures the variance due to departures from multiplicative fitness interactions among pairs of loci, is equal to

$$V_2 = n(n-1)\,\mathrm{E}[\varepsilon_{jk}^2 p_j q_j p_k q_k] \approx (n\bar{\mu})^2\,\mathrm{E}\left[\frac{\varepsilon_{jk}^2}{\tilde{a}_{j,\phi}\tilde{a}_{k,\phi}}\right], \quad (7)$$

where we have kept only the leading order term in ε_{jk} which has allowed us to assume that the allele frequencies are $q_j \approx \mu/\tilde{a}_{j,\phi}$ and $q_k \approx \mu/\tilde{a}_{k,\phi}$ where $\tilde{a}_{j,\phi}$ and $\tilde{a}_{k,\phi}$ measure the strength of selection against mutants at loci j and k (Barton, 1995, p. 134). Ignoring three-way and higher interactions, the strength of selection at the modifier locus can be written as

$$s_i \approx -\tfrac{1}{2}\overline{\delta r}(n\bar{\mu})^2 F, \quad (8)$$

where

$$F = \mathrm{E}[\varepsilon_{jk}]\,\mathrm{E}\left[\frac{1}{r_{ijk}r_{jk}}\left(\frac{1}{r_{ik}}+\frac{1}{r_{ij}}-1\right)\right] + \mathrm{E}\left[\frac{\varepsilon_{jk}^2}{\tilde{a}_{j,\phi}\tilde{a}_{k,\phi}}\right]\mathrm{E}\left[\frac{1}{r_{ijk}r_{jk}}\right].$$

For there to be selection for a modifier allele that increases recombination, F must be negative.

Variable epistatic interactions will always make F less negative by contributing to the second, strictly positive half of the equation for F. If the modifier is very tightly linked to the selected loci, this effect will be negligible. When the modifier is loosely linked, however, variable epistatic interactions will often cause F to be positive even when F would be negative in the absence of variability. In the case of an unlinked modifier, F will be negative only if

$$3\mathrm{E}[\varepsilon_{jk}] + \mathrm{E}\left[\frac{\varepsilon_{jk}^2}{\tilde{a}_{j,\phi}\tilde{a}_{k,\phi}}\right] < 0.$$

To proceed, we again standardize our measure of epistasis to the product of the selection coefficients, defining ε_{jk}^* as $\varepsilon_{jk}/(\tilde{a}_{j,\phi}\tilde{a}_{k,\phi})$ and rewriting the above condition as

$$3\mathrm{E}[\varepsilon_{jk}^*(\tilde{a}_{j,\phi}\tilde{a}_{k,\phi})] + \mathrm{E}[(\varepsilon_{jk}^*)^2(\tilde{a}_{j,\phi}\tilde{a}_{k,\phi})] < 0.$$

Assuming that there is little covariance between ε_{jk}^* and the selection coefficients $\tilde{a}_{j,\phi}$ and $\tilde{a}_{k,\phi}$ (which may or may not be biologically reasonable), this requires that

$$3\overline{\varepsilon^*} + \mathrm{Var}(\varepsilon^*) + (\overline{\varepsilon^*})^2 < 0, \quad (9)$$

where the mean and variance of ε_{jk}^* are given by $\overline{\varepsilon^*}$ and $\mathrm{Var}(\varepsilon^*)$, respectively. In the absence of variable epistatic interactions ($\mathrm{Var}(\varepsilon^*) = 0$), increased recombination will only be favored when $-3 < \overline{\varepsilon^*} < 0$, which is, to leading order, the same condition as obtained in the three-locus model, (3). Variability in the strength of interactions among loci makes condition (9) more difficult to satisfy. It will be impossible to satisfy if $\mathrm{Var}(\varepsilon^*)$ is greater than three times the magnitude of $\overline{\varepsilon^*}$.

We can convert these conditions into constraints on the genome-wide mutation rate as follows. From now on we focus on the fitness function (2) and begin by assuming that there is no variation in the strength of interactions (α and β are constant). Using Eq. (2) in Definition (4), the apparent strength of selection is approximately $\alpha + \beta\bar{n}$ when α and β are both small. Similarly, the apparent strength of epistasis is approximately $-\beta$ (also shown by Barton, 1995, Eqs. A3.2). Putting these together gives a value of $\varepsilon^* = -\beta/(\alpha+\beta\bar{n})^2$. For recombination to be favored if the modifier is unlinked, ε^* must be greater than -3 which requires that

$$\bar{n} > \frac{1}{\sqrt{3\beta}} - \frac{\alpha}{\beta}. \quad (10)$$

Equations (A3.3) and (A3.5) of Barton (1995) provide a relationship between the mean number of mutations and the mutation rate for this model:

$$U \approx (\alpha+\beta\bar{n})\left(\bar{n} - \frac{\beta\bar{n}^2}{2}\,\mathrm{E}\left[\frac{1}{r_{jk}}\right]\right) \quad (11)$$

Together Eqs. (10) and (11) indicate that, for increased recombination to be favored when there is negative epistasis, the genome wide mutation rate must satisfy

$$U > \frac{1}{3} - \frac{\alpha}{\sqrt{3\beta}} - \mathrm{E}\left[\frac{1}{r_{jk}}\right](\alpha - \sqrt{\beta/3})^2/(2\sqrt{3\beta}). \quad (12)$$

For instance, for selection to favor the maintenance of free recombination among all genes ($E[1/r_{jk}] = 2$) and using the estimates from Charlesworth (1990) of $\alpha = 0.002$ and $\beta = 0.0008$, the genome wide mutation rate must be larger than 0.288, in agreement with the critical value found by Charlesworth (1990). Interestingly, Eq. (12) indicates that, as long as log fitness is a quadratic function of the number of mutations carried in the genome, increased recombination will always be favored if there is synergistic epistasis and if $U > 1/3$, although as pointed out by Charlesworth (1990), the strength of selection in favor of recombination weakens as the genome map length increases.

Let us now consider the effect of variable epistatic interactions. Equation (9) can be rearranged as

$$\overline{\varepsilon^*} > -3 + \frac{\mathrm{Var}(\varepsilon^*)}{-\overline{\varepsilon^*}}. \quad (13)$$

Making the simplifying assumption that the variance in ε^* is some constant (c) times the mean value of $-\varepsilon^*$ (we are assuming here that ε^* is negative), Eq. (12) becomes

$$U > \frac{1}{3-c} - \frac{\alpha}{\sqrt{(3-c)\beta}} - E\left[\frac{1}{r_{jk}}\right](\alpha - \sqrt{\beta/(3-c)})^2/(2\sqrt{(3-c)\beta}). \quad (14)$$

Using the parameter values of Charlesworth (1990), free recombination among all loci will be favored if $U > 0.288$ when $c = 0$, $U > 0.442$ when $c = 1$, $U > 0.905$ when $c = 2$, and will never be favored if $c \geqslant 3$. The essential question to which we do not have an answer is whether the negative epistasis that has been observed experimentally (Mukai, 1969) is caused by numerous pairwise interactions of approximately the same strength (low $\mathrm{Var}(\varepsilon^*)$) or is due to a combination of strongly interacting mutations (that may, for example, be in the same biosynthetic pathway) and mutations with little interaction (high $\mathrm{Var}(\varepsilon^*)$). If interactions are highly variable ($\mathrm{Var}(\varepsilon^*) \gg |\overline{\varepsilon^*}|$), the evolution of increased recombination will be restricted and may only occur under fairly tight linkage among loci.

For partially linked modifiers, decreased recombination will still be favored whenever F in Eq. (8) for s_i is positive. Variance in epistatic interactions will always decrease the advantage of recombination, either selecting for decreased recombination rates or weakening the amount of selection in favor of increased recombination. Under tight linkage between the modifier and the loci under selection, however, increased recombination will generally be favored when there is negative epistasis.

CONCLUSIONS

This paper serves three main purposes: to make a qualitative comparison between the results of three-locus and multi-locus models, to correct the analysis of Feldman *et al.* (1980) (see appendix), and to point out that variability in the strength of epistatic interactions among loci makes the evolution of increased recombination more problematic.

A wide variety of models based on very different assumptions have been developed to analyse the evolution of recombination in the presence of deleterious mutations:

1. polygenic models which record only the number of mutations and not their nature and which assume a fitness function like (2) (Kondrashov, 1984; Charlesworth, 1990);

2. multi-locus models which assume that modifiers have weak effects and populations are near quasi-linkage equilibrium but which are otherwise fairly general (Barton, 1995); and

3. three-locus models which do not assume anything about fitness interactions or modifier effects, but which consider only two viability loci (appendix, this paper; Feldman *et al.* 1980).

These models share some general features leading us to the following picture of the evolution of recombination in the presence of deleterious mutations in large populations. If multiple mutations tend to dampen each others' effects (diminishing-returns or positive epistasis), decreased recombination rates are always favored. If, on the other hand, mutations tend to be more harmful than expected when combined (synergistic or negative epistasis), increased recombination rates are favored, but only under a restricted set of conditions. In all models, modifier alleles that increase recombination are favored if the modifier is tightly linked to the viability loci. *Decreased* recombination rates can, however, be favored under synergistic epistasis if the modifier locus is loosely linked. For example, in the three-locus model analysed in this paper, unlinked modifiers that increase recombination can only invade if epistasis is weak and negative (roughly $-s^2(3-s)/(1-s) < \varepsilon < 0$). Even for more tightly linked modifiers ($0.01 < R < 0.5$), recombination

is often favored only within a small range of ε as shown in Figs. 3 and 4.

In multi-locus formulations, particular functions are often used to describe the fitness interactions of multiple mutations, including quadratic functions, threshold functions and linear functions (Kondrashov, 1984; Charlesworth, 1990). These functions, however, obscure the fact that fitness interactions are likely to vary widely, with some loci interacting strongly and others weakly. Data, such as Mukai's (1969) data from *D. melanogaster*, describe the average interaction between random mutations; this also serves to mask the variability that occurs among the lines as new mutations arise which may or may not interact with previous mutations.

The extent of variability in epistasis is unknown, although there are several reasons to believe that it may be large. Mutations that affect the same metabolic pathway are more likely to interact strongly than mutations that affect different metabolic pathways. Similarly, mutations that occur within different components of a multimeric protein are more likely to interact than mutations within different proteins. Mutations that are expressed at the same time during development and/or within the same tissues may interact more strongly. Tightly linked mutations (e.g. mutations within the same gene) that together alter the secondary and/or tertiary structure of DNA (or mRNA or a protein) will have a stronger epistatic effect on fitness than unlinked mutations that do not interact in the determination of structure. Indeed, it seems much more plausible to assume that epistatic interactions are highly variable than the reverse.

In this paper, we have shown that it is critical that we account for this variability. Even if the epistasis falls, on average, within the range necessary for the evolution of increased recombination, variation among pairs of loci in the extent of epistasis means that some pairs of loci may be characterized by an ε that falls out of the range. If the variation in epistasis is very large, more loci will fall out of the range and favor decreased recombination than fall within the range. Adapting the analysis of Barton (1995), we have shown that variation in the extent of pair-wise epistasis restricts the conditions under which increased recombination can evolve with loosely linked loci. Defining the relative strength of epistasis as $\varepsilon^* = \varepsilon/s^2$, we found that an approximate condition that must be met for increased recombination to evolve is that $-3 + \mathrm{Var}(\varepsilon^*)/|\overline{\varepsilon^*}| < \overline{\varepsilon^*} < 0$ when the modifier is unlinked. Our analysis indicates that variable epistatic interactions may greatly restrict the degree of recombination that will evolve in a sexual system. In completely linked genomes, recombination will always be favored as long as there is negative epistasis, but high rates of recombination may be selected against, especially if there are highly variable epistatic interactions among mutants.

Previous analyses have concluded that increased recombination can be favored even at loosely linked modifier loci if there is quadratic selection with synergistic epistasis and if the genome-wide mutation rate is higher than about 1/3 (Kondrashov, 1984; Charlesworth, 1990; Barton, 1995). These results underestimate this critical mutation rate, since they do not account for variable epistatic relationships among loci. The qualitative conclusion that deleterious mutations can, under synergistic epistasis, favor the evolution of high rates of recombination even in large genomes may therefore be wrong. If the epistasis varies widely among different pairs of loci, if the genome-wide mutation rate is actually much lower than 1/3, or if there is diminishing-returns epistasis, then recombination levels across a genome will be selected to decrease. Further empirical measurements of these critical parameters will allow us to confirm or to reject the hypothesis that linkage relationships currently reflect the evolution of recombination in response to recurrent deleterious mutation.

APPENDIX

Recursions. Here we develop the recursion equations for a three-locus model in which selection acts upon adult haploid individuals according to their genotypes at two loci (A and B). Selection acts against these mutations, such that the viabilities of AB, Ab, aB and ab are 1, $W_1 = 1-s$, $W_1 = 1-s$, and $W_2 = (1-s)^2 + \varepsilon$, respectively. Sexual reproduction is assumed with the recombination rate between the selected loci controlled by a third locus (M, see Table 1). With two alleles at each of three loci, there are eight possible haploid chromosomes $\{MAB, MAb, MaB, Mab, mAB, mAb, maB, mab\}$, whose frequencies among adult haploids are $\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$. In developing the recursion equations, the order in which recombination and mutation proceeds is immaterial; both orderings lead to the same equations. Assume that recombination occurs first and let x_i^r denote the frequency of chromosome i after recombination:

$$\begin{aligned} x_1^r = x_1 &- r_1(x_1x_4 - x_2x_3) \\ &- (r_2 + R - 2r_2R)(x_1x_6 - x_2x_5) \\ &- R(x_1x_7 - x_3x_5) - (R - r_2R)(x_1x_8 - x_4x_5) \\ &+ r_2R(x_3x_6 - x_2x_7) \\ &+ r_2(x_2x_7 - x_1x_8) \end{aligned}$$

$$
\begin{aligned}
x_2^r &= x_2 + r_1(x_1x_4 - x_2x_3) \\
&\quad + (r_2 + R - 2r_2R)(x_1x_6 - x_2x_5) \\
&\quad - R(x_2x_8 - x_4x_6) + (R - r_2R)(x_3x_6 - x_2x_7) \\
&\quad - r_2R(x_1x_8 - x_4x_5) + r_2(x_1x_8 - x_2x_7) \\
x_3^r &= x_3 + r_1(x_1x_4 - x_2x_3) \\
&\quad - (r_2 + R - 2r_2R)(x_3x_8 - x_4x_7) \\
&\quad + R(x_1x_7 - x_3x_5) - (R - r_2R)(x_3x_6 - x_2x_7) \\
&\quad + r_2R(x_1x_8 - x_4x_5) - r_2(x_3x_6 - x_4x_5) \\
x_4^r &= x_4 - r_1(x_1x_4 - x_2x_3) \\
&\quad + (r_2 + R - 2r_2R)(x_3x_8 - x_4x_7) \\
&\quad + R(x_2x_8 - x_4x_6) + (R - r_2R)(x_1x_8 - x_4x_5) \\
&\quad - r_2R(x_3x_6 - x_2x_7) - r_2(x_4x_5 - x_3x_6) \\
x_5^r &= x_5 - r_3(x_5x_8 - x_6x_7) \\
&\quad + (r_2 + R - 2r_2R)(x_1x_6 - x_2x_5) \\
&\quad + R(x_1x_7 - x_3x_5) + (R - r_2R)(x_1x_8 - x_4x_5) \\
&\quad - r_2R(x_3x_6 - x_2x_7) - r_2(x_4x_5 - x_3x_6) \\
x_6^r &= x_6 + r_3(x_5x_8 - x_6x_7) \\
&\quad - (r_2 + R - 2r_2R)(x_1x_6 - x_2x_5) \\
&\quad + R(x_2x_8 - x_4x_6) - (R - r_2R)(x_3x_6 - x_2x_7) \\
&\quad + r_2R(x_1x_8 - x_4x_5) - r_2(x_3x_6 - x_4x_5) \\
x_7^r &= x_7 + r_3(x_5x_8 - x_6x_7) \\
&\quad + (r_2 + R - 2r_2R)(x_3x_8 - x_4x_7) \\
&\quad - R(x_1x_7 - x_3x_5) + (R - r_2R)(x_3x_6 - x_2x_7) \\
&\quad - r_2R(x_1x_8 - x_4x_5) + r_2(x_1x_8 - x_2x_7) \\
x_8^r &= x_8 - r_3(x_5x_8 - x_6x_7) \\
&\quad - (r_2 + R - 2r_2R)(x_3x_8 - x_4x_7) \\
&\quad - R(x_2x_8 - x_4x_6) - (R - r_2R)(x_1x_8 - x_4x_5) \\
&\quad + r_2R(x_3x_6 - x_2x_7) + r_2(x_2x_7 - x_1x_8).
\end{aligned}
$$

After mutation and selection, the chromosome frequencies in the next generation, x_i', become

$$
\begin{aligned}
\bar{W}x_1' &= (1-\mu)^2 x_1^r \\
\bar{W}x_2' &= W_1(\mu(1-\mu)\, x_1^r + (1-\mu)\, x_2^r) \\
\bar{W}x_3' &= W_1(\mu(1-\mu)\, x_1^r + (1-\mu)\, x_3^r) \\
\bar{W}x_4' &= W_2(\mu^2 x_1^r + \mu x_2^r + \mu x_3^r + x_4^r) \\
\bar{W}x_5' &= (1-\mu)^2\, x_5^r \\
\bar{W}x_6' &= W_1(\mu(1-\mu)\, x_5^r + (1-\mu)\, x_6^r) \\
\bar{W}x_7' &= W_1(\mu(1-\mu)\, x_5^r + (1-\mu)\, x_7^r) \\
\bar{W}x_8' &= W_2(\mu^2 x_5^r + \mu x_6^r + \mu x_7^r + x_8^r),
\end{aligned}
$$

where $\bar{W}$ is the mean fitness and is equal to the sum of the right hand sides of these equations.

Stability analysis. With the M allele fixed, the above system equilibrates at a mutation-selection balance if selection is strong relative to the mutation rate. Specifically, when $r_1 = 0$, the existence and stability of a mutation-selection balance requires that $1-\mu > W_1$ and $(1-\mu)\, W_1 > W_2$ (Karlin & McGregor, 1972; Feldman *et al.*, 1980), conditions that will be assumed to hold throughout the following analysis. Equilibrium values of each variable are denoted as $\hat{x}_i$ and $\hat{\bar{W}}$. In addition, the disequilibrium at the mutation-selection balance is defined as $\hat{D} = \hat{x}_1\hat{x}_4 - \hat{x}_2\hat{x}_3$. Feldman *et al.* (1980) showed that the sign of $\hat{D}$ is determined by the nature of the epistasis between the two loci and has the same sign as $W_2 - W_1^2 = \varepsilon$. In addition, they showed that $\hat{\bar{W}} > W_1(1-\mu)$, which will be important in evaluating the subsequent equations. The local stability of the equilibrium with M fixed to the introduction of a new modifier allele, m, was studied using the linearized recursions for the haplotypes bearing m. In matrix form, these linearized recursions constitute a 4×4 matrix with strictly positive elements as long as $R > 0$ (Feldman *et al.*, 1980). Therefore, by the Perron–Frobenius theorem (Gantmacher, 1989), the eigenvalue of greatest magnitude (λ_L) will be positive. If this leading eigenvalue is greater than one, m will tend to rise in frequency and invade the population. Details of the stability analysis are available upon request or may be accessed on the World Wide Web at http://www.zoology.ubc.ca/~otto/Research/Recombination/Math.html. Since the characteristic polynomial, $f_H(\lambda)$, of the matrix has a positive λ^4 term, a sufficient condition for its leading eigenvalue, λ_L, to be greater than one is for $f_H(\lambda)$ to be negative at $\lambda = 1$. Although this is not a necessary condition, numerical evaluations of the characteristic polynomial failed to find any instances where $f_H(1) > 0$ and yet the leading eigenvalue was greater than one. We therefore assume that the equilibrium with M fixed will be stable if $f_H(1)$ is positive, whereas invasion will occur if $f_H(1)$ is negative. $f_H(1)$ is not equal to Eq. (7) of Feldman *et al.* (1980), but rather equals

$$f_H(1) = \frac{(r_2 - r_1)\,\hat{D}(1-\mu)^2}{\hat{\bar{W}}^4 \hat{x}_1} \times \{[\hat{\bar{W}} - W_1(1-\mu)]\,K_1 - R[2(\hat{\bar{W}} - W_1(1-\mu))(1-r_2)(1-\mu) \times K_2 + r_2(1-\mu)^2 K_3] - R^2[(1-\mu)^2(1-2r_2)K_3]\}, \quad (15)$$

where

$$K_1 = (\hat{\bar{W}} - W_1(1-\mu))(\hat{\bar{W}} - W_2) - r_2(1-\mu)K_2,$$
$$K_2 = -W_1(\hat{\bar{W}} - W_2) + \hat{\bar{W}}(W_1 - W_2)(\hat{x}_1 + \hat{x}_2),$$
$$K_3 = -W_1^2(\hat{\bar{W}} - W_2) + 2\hat{\bar{W}}W_1(W_1 - W_2)(\hat{x}_1 + \hat{x}_2).$$

Equation (15) differs from Eq. (7) of Feldman *et al.* (1980) by the factor that multiplies R. Direct numerical evaluations of the stability matrix are consistent with this revised equation for the value of $f_H(1)$.

Weak epistasis. With very weak epistasis ($W_2 - W_1^2 = \varepsilon \ll s, \mu$), one can obtain approximations for $\hat{\bar{W}}$, $\hat{D}$, and $\hat{p}_a = \hat{p}_b = 1 - (\hat{x}_1 + \hat{x}_2)$. To order ε, these become

$$\hat{\bar{W}} = (1-\mu)^2 - \varepsilon \frac{(1-\mu)^2 \mu^2 r_1}{(1-W_1)^2(1 - W_1^2 + r_1 W_1^2)} + O(\varepsilon^2)$$

$$\hat{D} = \varepsilon \frac{\mu^2(1-\mu-W_1)^2}{(1-\mu)^2(1-W_1)^4(1-W_1^2+r_1W_1^2)} + O(\varepsilon^2)$$

$$\hat{p}_a = \frac{\mu W_1}{1-\mu-W_1+\mu W_1} + \varepsilon \frac{\mu^2(1-\mu-W_1+r_1W_1)}{(1-\mu)(1-W_1)^3(1-W_1^2+r_1W_1^2)} + O(\varepsilon^2)$$

Using these approximations and $x_1 = \hat{p}_a\hat{p}_b + \hat{D}$ in Eq. (15) leads to

$$f_H(1) = \varepsilon\mu^2(r_2 - r_1)(1-\mu-W_1)^2 \times \frac{\left(\begin{array}{l}[(1-\mu)^2 - W_1^2(1-r_2) \\ \quad + RW_1^2(2-3r_2) - R^2W_1^2(1-2r_2)]\end{array}\right)}{(1-\mu)^4(1-W_1)^2(1-W_1^2+r_1W_1^2)} + O(\varepsilon^2). \quad (16)$$

At $R = 0$, the term in square brackets is positive by the assumption that $1 - \mu > W_1$. Because of the $-R^2$ term, this term will eventually become negative for large enough R, but it is still positive at $R = \frac{1}{2}$. That is, there is a critical value of the recombination rate (R^*) above which the stability conditions would switch, but this critical value is always above $\frac{1}{2}$ when epistasis is very weak. Therefore, for all biologically reasonable values of R, the value of $f_H(1)$ depends simply on the sign of $(r_2 - r_1)\,\varepsilon$. Consequently, invasion of the m allele will occur when it increases the recombination rate and there is negative (synergistic) epistasis or when it decreases the recombination rate and there is positive (diminishing-returns) epistasis.

General case. As noted by Feldman *et al.* (1980), since K_1 is positive under the assumptions of the model, $f_H(1)$ has the sign of $(r_2 - r_1)\,\varepsilon$ when $R = 0$ and the m allele will invade if $(r_2 - r_1)\,\varepsilon < 0$. When the modifier is tightly linked to the loci under selection, increased recombination is favored under negative epistasis while decreased recombination is favored under positive epistasis as was found when very weak epistasis was assumed.

As written in (15), $f_H(1)$ is a quadratic function of R. For very weak epistasis, the quadratic term of this function has the sign of $-(r_2 - r_1)\,\varepsilon$ (opposite to the sign of $f_H(1)$ at $R = 0$). Consequently, as we increase R from zero, $f_H(1)$ must eventually change sign once and only once (at a point, R^*). This R^* is always greater than $\frac{1}{2}$ with very weak epistasis, but may fall within $[0, \frac{1}{2}]$ for stronger epistasis. We can determine whether there will be a switch in the stability of the system to invasion by a modifier for biologically reasonable values of R by determining whether $f_H(1)$ at $R = \frac{1}{2}$ still has the same sign as $f_H(1)$ at $R = 0$. There is an additional complication with strong epistasis, however, since the quadratic term of $f_H(1)$ may sometimes have the sign of $(r_2 - r_1)\,\varepsilon$ which would make it possible for there to be two positive roots of $f_H(1)$ between $R = 0$ and $R = \frac{1}{2}$. A numerical search did not, however, reveal any cases where $f_H(1)$ changed sign more than once for R within $[0, \frac{1}{2}]$. We assume, therefore, that either the largest eigenvalue passes through 1 once and only once as R goes from 0 to $\frac{1}{2}$, or that it remains on the same side of 1 for all $0 \leqslant R \leqslant \frac{1}{2}$. To determine whether a switch in stability can occur for biologically meaningful values of R, the sign of Eq. (15) at $R = \frac{1}{2}$ may be compared to that at $R = 0$. Using equilibrium equations that relate $\hat{x}_1$ and μ to $\hat{\bar{W}}$ and $(\hat{x}_1 + \hat{x}_2)$, $f_H(1)$ at $R = \frac{1}{2}$ is found to equal

$$\frac{\left(\begin{array}{l}(r_2 - r_1)\,\hat{D}(1-\mu)^2(2\hat{\bar{W}} - (1-\mu)\,W_1) \\ \quad [(1 - 2W_1 + W_2)\hat{\bar{W}}(\hat{x}_1 + \hat{x}_2) + (1 - W_1)(\hat{\bar{W}} - W_2)]\end{array}\right)}{4\hat{\bar{W}}^3(\hat{x}_1 + \hat{x}_2)}. \quad (17)$$

When the term in square brackets is positive, $f_H(1)$ at $R=0$ and at $R=\frac{1}{2}$ have the same sign and no switch in stability is observed. The term in square brackets is positive when epistasis is positive ($W_2 > W_1^2$). The term in square brackets is also positive if there is negative *but weak* epistasis on a multiplicative scale such that $1-2W_1+W_2$ is still positive (i.e. there is still positive epistasis on an additive scale). If there is strong negative epistasis, however, the first term in the square brackets can be sufficiently negative for a switch in stability to be observed between $R=0$ and $R=\frac{1}{2}$. In this case, even though epistasis is negative, only modifier alleles that *decrease* recombination can invade when $R=\frac{1}{2}$. To determine when the term in square brackets can be negative, assume that mutation is weak relative to the selection coefficients. Then, to leading order in the mutation rate, the term in square brackets becomes $2-3W_1+W_1W_2+O(\mu)$ (since $\hat{\bar{W}}=1-O(\mu)$ and $\hat{x}_1+\hat{x}_2=1-O(\mu)$), which will be negative whenever

$$W_1 < 3-\frac{2}{W_1} \tag{18}$$

or

$$\varepsilon < \frac{-s^2(3-s)}{(1-s)}, \tag{19}$$

in which case only modifiers that decrease recombination can invade. The implication of this analysis is that loosely linked modifiers that increase the rate of recombination are only able to invade for a very limited parameter range, that is, only when

$$\frac{-s^2(3-s)}{(1-s)} < \varepsilon < 0.$$

The breadth of this "window" in which increased recombination is favored at loosely linked modifier loci is small, extremely so when selection is weak (when the breadth approaches $3s^2$).

Diploid model. Feldman *et al.* (1980) also examined the above model when selection acts upon the diploid stage rather than the haploid stage. The viability of a diploid individual is given by W_{ij}, where i and j denote the two component haplotypes of the diploid ($i, j = 1, 2, 3, 4$ for AB, Ab, aB, and ab, respectively). Making the assumptions of Feldman *et al.* (1980) that $W_{ij}=W_{ji}$, $W_{2j}=W_{3j}$ and $W_{j2}=W_{j3}$ and defining the marginal fitness of an allele j as $W_{j\cdot}=\sum_{i=1}^{4} W_{1i}x_i$, we find that the value of the characteristic polynomial evaluated at $\lambda=1$ equals

$$\begin{aligned} f_H(1) = {} & \frac{(r_2-r_1)(W_{14}\hat{x}_1\hat{x}_4 - W_{23}\hat{x}_2\hat{x}_3)(1-\mu)^2}{\hat{x}_1\hat{\bar{W}}^4} \\ & \times \{[\hat{\bar{W}} - W_{2\cdot}(1-\mu)]\, L_1 \\ & - R[2(\hat{\bar{W}} - W_{2\cdot}(1-\mu))(1-r_2)(1-\mu) \\ & \times L_2 + r_2(1-\mu)^2 L_3] \\ & - R^2[(1-\mu)^2(1-2r_2) L_3]\}, \end{aligned} \tag{20}$$

where

$$\begin{aligned} L_1 = {} & (\hat{\bar{W}} - (1-\mu)\, W_{2\cdot})(\hat{\bar{W}} - W_{4\cdot}) - r_2(1-\mu) L_2 \\ L_2 = {} & -(\hat{\bar{W}} - W_{2\cdot})(\hat{x}_2 W_{24} + \hat{x}_1 W_{14}) \\ & -(\hat{\bar{W}} - W_{4\cdot})(\hat{x}_2 W_{22} + \hat{x}_4 W_{24}) \\ L_3 = {} & -2W_{2\cdot}(\hat{\bar{W}} - W_{2\cdot})(\hat{x}_1 W_{14} + \hat{x}_2 W_{24}) \\ & + W_{2\cdot}(\hat{\bar{W}} - W_{4\cdot})(\hat{x}_1 W_{12} - \hat{x}_4 W_{24}). \end{aligned}$$

At $R=0$, this reduces to Eq. (11) of Feldman *et al.* (1980) under the assumption that there are no cis-trans effects on viability, such that an AB/ab genotype has the same fitness as an Ab/aB genotype ($W_{14}=W_{23}$; but see Nordborg *et al.* 1995).

At $R=0$, as observed by Feldman *et al.* (1980), invasion of a modifier allele that increases recombination will occur only if $\hat{D}<0$, while invasion of a modifier allele that decreases recombination will occur if $\hat{D}>0$. At $R=1/2$, Eq. (20) becomes

$$\begin{aligned} & \frac{(r_2-r_1)(W_{14}\hat{x}_1\hat{x}_4 - W_{23}\hat{x}_2\hat{x}_3)(1-\mu)^2}{\hat{x}_1\hat{\bar{W}}^4} \\ & \times \frac{2\hat{\bar{W}} - W_{2\cdot}(1-\mu)}{4} \\ & \times [2(\hat{\bar{W}} - W_{2\cdot}(1-\mu))(\hat{\bar{W}} - W_{4\cdot}) - (1-\mu) L_3/W_{2\cdot}]. \end{aligned}$$

Again, a switch in stability will occur between $R=0$ and $R=1/2$ only if the term in square brackets is negative. To leading order in μ, this term equals $2-3W_{12}+W_{12}W_{14}+O(\mu)$. Defining $W_{12}=1-s$ and $W_{14}=(1-s)^2+\varepsilon$, we regain condition (19). Thus, when mutation is weak relative to selection, a switch in stability occurs in the diploid model under the same condition as in the haploid model: there must be relatively strong negative epistasis.

Final note. The above analysis is consistent with the qualitative conclusions of Feldman *et al.* (1980) up to a point. They argued that a tightly linked modifier would invade if it increased recombination and $\varepsilon < 0$ or if it decreased recombination and $\varepsilon > 0$, "but if the linkage is loose enough (when R^* exists), it may be eliminated." Numerically evaluating their stability condition indicates that an R^* between 0 and $\frac{1}{2}$ can exist (a switch in stability may occur) for both positive and negative epistasis. In contrast, this reanalysis indicates that an R^* between 0 and $\frac{1}{2}$ can only exist when epistasis is negative. Modifier alleles that increase recombination are *never* favored if there is positive epistasis and are favored under negative epistasis only if the modifier is sufficiently tightly linked.

ACKNOWLEDGEMENTS

We thank Nick Barton for many insightful discussions about the evolution of recombination and for suggesting the use of Eq. (5) to examine variable epistatic interactions. We also thank Nick Barton, Brian Charlesworth, Michael Whitlock and an anonymous reviewer for helpful comments on the manuscript. S.P.O. was sponsored by a Science and Engineering Research Council grant to Nick Barton and by a grant from the Natural Sciences and Engineering Research Council of Canada. M.W.F. was sponsored by NIH Grant GM28016.

REFERENCES

Barton, N. H. 1995. A general model for the evolution of recombination, *Genet. Res. Cambridge* **65**, 123–144.

Bergman, A., Otto, S. P., and Feldman, M. 1995. On the evolution of recombination in haploids and diploids. I. Deterministic models, *Complexity* **1**, 49–57.

Brooks, L. 1988. The evolution of recombination rates, *in* "The Evolution of Sex: An Examination of Current Ideas" (R. E. Michod and B. R. Levin, Eds.), pp. 87–105, Sinauer, Sunderland, MA.

Charlesworth, B. 1976. Recombination modification in a fluctuating environment, *Theor. Popul. Biol.* **83**, 181–195.

Charlesworth, B. 1990. Mutation selection balance and the evolutionary advantage of sex and recombination, *Genet. Res. Cambridge* **55**, 199–221.

Charlesworth, B. 1993. The evolution of sex and recombination in a varying environment, *J. Heredity* **84**, 345–350.

Charlesworth, B., and Barton, N. 1996. Recombination load associated with selection for increased recombination, *Genet. Res. Cambridge* **67**, 27–41.

Charlesworth, B., Charlesworth, D., and Morgan, M. 1990. Genetic loads and estimates of mutation rates in highly inbred plant populations, *Nature* **347**, 380–382.

Crow, J. F. 1970. Genetic loads and the cost of natural selection, *in* "Mathematical Topics in Population Genetics" (K.-I. Kojima, Ed.), pp. 128–177, Springer-Verlag, New York.

Eshel, I., and Feldman, M. W. 1970. On the evolutionary effect of recombination, *Theor. Popul. Biol.* **1**, 88–100.

Feldman, M. W. 1972. Selection for linkage modification. I. Random mating populations, *Theor. Popul. Biol.* **3**, 324–346.

Feldman, M. W., Christiansen, F. B., and Brooks, L. D. 1980. Evolution of recombination in a constant environment, *Proc. Nat. Acad. Sci. USA* **77**, 4838–4841.

Feldman, M. W., Otto, S. P., and Christiansen, F. B. 1997. Population genetic perspectives on the evolution of recombination, *Ann. Rev. Genet.* **30**, 261–295.

Fisher, R. A. 1930. "The Genetical Theory of Natural Selection," Oxford Univ. Press, Oxford.

Gantmacher, F. R. 1989. "The Theory of Matrices," Vol. 2, Chelsea, New York.

Hamilton, W. D. 1980. Sex versus non-sex versus parasite, *OIKOS* **35**, 282–290.

Houle, D., Hughes, K. A., Hoffmaster, D. K., Ihara, J. *et al.* 1994. The effects of spontaneous mutation on quantitative traits. I. Variances and covariances of life history traits, *Genetics* **138**, 773–785.

Karlin, S., and McGregor, J. 1972. Application of method of small parameters to multi-niche population genetic models, *Theor. Popul. Biol.* **3**, 186–209.

Keightley, P. D. 1994. The distribution of mutation effects in *Drosophila melanogaster*, *Genetics* **138**, 1315–1322.

Kibota, T. T. and Lynch, M. 1996. Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*, *Nature* **381**, 694–696.

Kimura, M. and Maruyama, T. 1966. The mutational load with epistatic gene interactions in fitness, *Genetics* **54**, 1303–1312.

Kondrashov, A. S. 1982. Selection against harmful mutations in large sexual and asexual populations, *Genet. Res. Cambridge* **40**, 325–332.

Kondrashov, A. S. 1984. Deleterious mutations as an evolutionary factor. I. The advantage of recombination, *Genet. Res. Cambridge* **44**, 199–217.

Kondrashov, A. S. 1988. Deleterious mutations and the evolution of sexual reproduction, *Nature* **336**, 435–440.

Kondrashov, A. S. 1993. Classification of hypotheses on the advantage of amphimixis, *J. Heredity* **84**, 372–387.

Kondrashov, A. S., and Crow, J. 1993. A molecular approach to estimating the human deleterious mutation rate, *Hum. Mutat.* **2**, 229–234.

Maynard–Smith, J. 1968. Evolution in sexual and asexual populations, *Am. Nat.* **102**, 469–473.

Michod, R., and Levin, B. 1988. "The Evolution of Sex: An Examination of Current Ideas," Sinauer, Sunderland, MA.

Mukai, T. 1964. The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability, *Genetics* **50**, 1–19.

Mukai, T. 1969. The genetic structure of natural populations of *Drosophila melanogaster*. VII. Synergistic interaction of spontaneous mutant polygenes controlling viability, *Genetics* **61**, 749–761.

Mukai, T., Chigusa, S. I., Mettler, L. E., and Crow, J. F. 1972. Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*, *Genetics* **72**, 335–355.

Muller, H. J. 1932. Some genetic aspects of sex, *Am. Nat.* **66**, 118–138.

Muller, H. J. 1964. The relation of recombination to mutational advance, *Mutat. Res.* **1**, 2–9.

Nei, M. 1967. Modification of linkage intensity by natural selection, *Genetics* **57**, 625–641.

Nordborg, M., Franklin, I., and Feldman, M. 1995. Effects of *cis–trans* viability selection on some two-locus models, *Theor. Popul. Biol.* **47**, 365–392.

Sturtevant, A., and Mather, K. 1938. The interrelations of inversions, heterosis, and recombination, *Am. Nat.* **72**, 447–452.

Whitlock, M., Phillips, P., Moore, F. B.-G., and Tonsor, S. J. 1995. Multiple fitness peaks and epistasis, *Annu. Rev. Ecol. Syst.* **26**, 601–629.