SOME ADVANTAGES AND DISADVANTAGES OF RECOMBINATION

SARAH P. OTTO

Department of Integrative Biology
UC Berkeley, Berkeley CA 94720

MARCUS W. FELDMAN

Department of Biological Sciences
Stanford, CA 94305

FREDDY B. CHRISTIANSEN

Department of Ecology and Genetics
University of Aarhus
DK-800 Aarhus C, Denmark

Abstract: Theory for the evolutionary properties of recombination is reviewed, both in terms of modifier genes and approaches that use group selection. Analytical and numerical methods are used to investigate the time to first appearance of an advantageous two locus chromosome. On average, whether or not a double mutant chromosome is advantageous, it will appear for the first time faster with recombination than without. A comparison is made with recent work on genetic algorithms. It is shown, using simple examples, that whether recombination speeds the production of a favorable multilocus chromosome depends on the shape of the fitness function and the initial allele frequencies.

INTRODUCTION

Recombination, including chromosomal segregation, shuffles together the genetic material carried by different members of a sexual species. This genetic mixing unties the evolutionary fate of alleles at one locus from the fate of alleles at neighboring loci and can increase the amount of genetic variation found within a population. In the process, however, recombination separates advantageous gene combinations, the very gene combinations that enabled the parents to survive and reproduce. Whether or not the adaptation of a population to an environment is more rapid in the presence of recombination, that is, whether or not recombination speeds up the evolutionary process, depends critically on the ways in which this process is modelled. As we shall

see, the effect of recombination depends on the population size, the initial population composition, and the selection regime under consideration.

Fisher, in 1930, and Muller, in 1932, both argued that recombination allows favorable mutants that arise in separate individuals to be combined within the same genome. This advantage of recombination rests upon three important assumptions. The first is that favorable mutations do not fix more rapidly than they arise, for, if they were to fix more rapidly, then advantageous mutations would not segregate simultaneously at different loci and recombination would not matter. The second is that favorable mutations are much less likely to co-occur within an individual through mutation than through the action of recombination. Finally, for the full impact of recombination to be assessed, the disadvantage of separating favorable gene combinations must be weighed along with the advantage of combining separate mutations. It is not obvious that these assumptions are valid. For instance, high rates of mutation favor concurrent segregation (the first assumption) but they also favor the production of double mutants through mutation (making the second assumption less tenable). We briefly discuss some important studies in which this issue has been quantitatively addressed.

Maynard-Smith (1968) followed the spread of advantageous mutations in a two locus model, assuming discrete generations, a large population, and multiplicative selection. He showed that genetic associations (linkage disequilibria) would not develop in the population if there were no such associations to begin with. Recombination, however, only affects the dynamics of a population if there is disequilibrium (see equation 1). Consequently, the presence or absence of recombination was irrelevant to the evolution of such a population. These results were extended to more general fitness schemes by Eshel and Feldman (1970). They considered fitness interactions (i.e. epistasis) between the two loci, as outlined in Table 1. If the fitness of the advantageous double mutant is greater than the product of the fitnesses of the single mutants (i.e. supermultiplicative selection), more of the favorable double mutants would be found in the population without recombination than with recombination. Essentially, in the supermultiplicative case, more double mutants are present than expected (positive linkage disequilibrium) and recombination only serves to separate the advantageous double mutant combination. With submultiplicative fitnesses, the number of favorable double mutants increases with recombination (see also Karlin (1973)). The sign of the linkage disequilibrium is critical here: when fitnesses are supermultiplicative, it becomes positive, while when they are submultiplicative, it becomes negative. Thus, in an infinite population, the advantage of recombination is sensitive to the fitness regime.

The preceding arguments have been made in the framework of large populations in

Table 1: Fitness regime. At the two-loci, the mutant alleles (a and b) are advantageous ($\sigma_1, \sigma_2, \sigma_3 > 0$) over the non-mutant alleles (A and B). The amount of epistasis is measured by γ ($\gamma = \sigma_1 - \sigma_2\sigma_3$).

	AB	Ab	aB	ab		Epistasis
Supermultiplicative	1	σ_3	σ_2	σ_1	($\sigma_1 > \sigma_2\sigma_3$)	Positive
Multiplicative	1	σ_3	σ_2	σ_1	($\sigma_1 = \sigma_2\sigma_3$)	Absent
Submultiplicative	1	σ_3	σ_2	σ_1	($\sigma_1 < \sigma_2\sigma_3$)	Negative

the absence of genetic drift. The Fisher-Muller hypotheses are, however, more appropriately addressed by finite population size models (Felsenstein, 1988). For instance, stochastic effects are important in Muller's (1964) proposal that, in the presence of mutation to deleterious alleles, recombination would permit the reconstitution of mutation-free individuals. In an infinite population, the mutation-free class would never entirely disappear. It is only in a finite population that this class would be lost by drift and could then be reformed by recombination (Haigh, 1978; Felsenstein, 1988).

Finite population size has been included in a series of quantitative studies examining the advantages of recombination. Crow and Kimura (1965) and Bodmer (1970), used $1/N$ as the frequency at which a (haploid) chromosome first appears, where N is the population size, but these studies assumed that sampling error (random genetic drift) was negligible. Crow and Kimura (1965) argued that in a population with no recombination only one lineage would eventually survive, and only those mutations that occurred in this lineage could possibly increase to fixation. With recombination, however, all mutations could potentially be incorporated. Using several approximations, they found that the relative rate of incorporation of new advantageous mutants may be several orders of magnitude larger with recombination, especially if the population size is large. Bodmer (1970), on the other hand, argued that the time until the most fit chromosome first appeared in the population was shorter with recombination and that this effect is larger with *small* population sizes. These results are difficult to evaluate, however, because the stochastic effects of genetic drift are ignored in both.

Other studies have assumed a finite population of constant size in every generation and have examined the effects of sampling. Hill and Robertson (1966) simulated the evolution of a finite population assuming a two-locus model with either additive selection (which, in terms of Table 1, is submultiplicative) or multiplicative selection. In both cases, they found that selection at one locus reduced the chance of fixation at the second locus, especially with tight linkage. Increased linkage also increased

3

the time required for the population to fix. These authors argued that the effective population size for one locus is reduced by selection on the second locus to the fraction of the population which carries the favorable allele at the second locus. Although not a part of their argument, the fact that negative disequilibrium generally developed in their simulations meant that the frequency of favorable alleles at the first locus was lower in that fraction of the population that also carried the favorable allele at the second locus than in the remaining fraction of the population. Recombination acts to reduce the linkage disequilibrium between a pair of linked genes reducing the amount that each interferes with the other's response to selection. This is called the Hill-Robertson effect.

Assuming equal fitnesses and directional mutation, Karlin (1973) showed analytically for a population of size two (!) that the time until first appearance of the double mutant was shorter with recombination, but the time until fixation was longer. A closer inspection of the time until fixation shows, however, that the dependence of the fixation time on recombination is very small (on the order of the mutation rate). Nevertheless, this study suggests that recombination in the absence of selection might increase the time required to fix mutant chromosomes, a result opposite to that found by Hill and Robertson (1966) with selection.

Felsenstein (1974) and Felsenstein and Yokoyama (1976) carried out simulation studies of a finite population under multiplicative selection. They confirmed the findings of Hill and Robertson (1966) and in the second paper extended the results by including a third locus (a modifier locus) that controlled the rate of recombination between the two-loci under selection. They found that an allele for free recombination significantly outcompeted an allele that eliminated recombination. They attributed this advantage of recombination to its effect of breaking down the linkage disequilibrium generated by random genetic drift in the presence of selection.

Recent efforts to use *genetic algorithms* to solve complicated optimization problems have generated a great deal of interest among computer scientists in the evolution of recombination in small populations. Genetic algorithms are procedures invented by Holland (see e.g. Holland, 1992) that attempt to find the optima of complicated functions using programs that simulate Darwinian natural selection. Potential solutions are treated as strings of symbols that are subject to mutation and may be recombined with one another to form new strings. The new strings are subject to selection on the basis of their performance at the task set by the user (Goldberg, 1989; Belew and Booker, 1991). The role of recombination in these simulations is the subject of considerable debate, remarkably analogous to the discussions about recombination in population genetics (e.g. Sumida et al., 1990; Schaffer and Eshelman, 1991).

We have performed a large simulation study that investigates the advantages and disadvantages of recombination in finite populations with different two-locus selective schemes. We find that recombination always hastens the first appearance of a double mutant, but we argue that measuring the time to first appearance is tantamount to measuring the advantages of recombination when recombination cannot be disadvantageous, that is, when there are no double mutants that may be separated by recombination. Recombination may either increase or decrease the time to fixation of the advantageous type depending on the selection regime. We suggest that these results are best understood in terms of the development of disequilibria and not, as argued by Hill and Robertson (1966), in terms of the effective population size.

METHODS

Consider a finite haploid population of size N. Two loci are both subject to recurrent mutations between the resident alleles (A and B) and the new alleles (a and b), with recombination between the two loci at a rate, r. Fitnesses are either equivalent for all genotypes (neutral) or are specified according to Table 1. We make the further simplifying assumption that selection is equivalent at the two loci ($\sigma_3 = \sigma_2$).

We assume that the finite haploid population produces an effectively infinite number of gametes and that these gametes unite at random to produce diploid zygotes which immediately undergo meiosis to produce haploid offspring. Selection then acts to change the proportion of the different offspring genotypes. If $x_1, x_2, x_3,$ and x_4 are the frequencies of the adult genotypes $ab, Ab, aB,$ and AB then the offspring distribution after selection (x_i^o) will take the form

$$
\begin{aligned}
W x_1^o &= \sigma_1\{(1-\mu_B)(1-\mu_B)x_1 + \mu_F(1-\mu_B)x_2 + \mu_F(1-\mu_B)x_3 + \mu_F^2 x_4 \\
&\qquad\qquad -rD(1-\mu_F-\mu_B)^2\} \\
W x_2^o &= \sigma_2\{(1-\mu_F)(1-\mu_B)x_2 + \mu_B(1-\mu_B)x_1 + \mu_F\mu_B x_3 + \mu_F(1-\mu_F)x_4 \\
&\qquad\qquad +rD(1-\mu_F-\mu_B)^2\} \qquad\qquad (1) \\
W x_3^o &= \sigma_2\{(1-\mu_F)(1-\mu_B)x_3 + \mu_B(1-\mu_B)x_1 + \mu_F\mu_B x_2 + \mu_F(1-\mu_F)x_4 \\
&\qquad\qquad +rD(1-\mu_F-\mu_B)^2\} \\
W x_4^o &= \{(1-\mu_F)(1-\mu_F)x_4 + \mu_B(1-\mu_F)x_2 + \mu_B(1-\mu_F)x_3 + \mu_B^2 x_1 \\
&\qquad\qquad -rD(1-\mu_F-\mu_B)^2\}
\end{aligned}
$$

where μ_F and μ_B are the forward and backward mutation rates, respectively, for both loci, D is the linkage disequilibrium ($D = x_1 x_4 - x_2 x_3$), and W is the sum of the right sides. In our simulations, we either ignore back-mutations ($\mu_B = 0$) or set the forward and backward mutation rates equal to one another ($\mu_F = \mu_B$). Finally, the next generation of N adults is chosen by random sampling from the offspring

distribution described by equation (1). Notice that selection acts on the haploid juvenile stage in this life cycle, and therefore affects the probability that a particular offspring may become established as one of the N sampled adults. This sampling regime is known as the Wright-Fisher multinomial sampling model.

All simulations were started with a population of N adults carrying the AB chromosome ($x_4 = 1$). The offspring distribution was then produced according to (1) and Wright-Fisher sampling performed to create the next generation. This procedure was repeated until fixation of the ab chromosome first occurred. The number of generations required until the first appearance of the ab chromosome within the adult population (T_A) and the number of generations to fixation (T_F) were both recorded. Note that with backward mutations, a population never permanently fixes upon ab, but will fix temporarily if the mutation rate is sufficiently small ($N\mu < 1$).

RESULTS

No Selection

Table 2 reports results from simulations without selection. Here, we are only interested in the case of unidirectional mutation, so that there is a definite increase in the number of mutant alleles. As can be seen, the time until first appearance of the double mutant always decreases with increasing recombination rate. On the other hand, the time to fixation shows no clear dependence on the recombination rate. The analysis of Karlin (1973), however, can be used to show that the time to fixation should increase with recombination, but that $\partial T_F / \partial r$ is on the order of the mutation rate and decreases in magnitude with r (for $N = 2$). In order to improve the accuracy of the results, we repeated the simulations with $N = 100$ and $\mu = 0.01$ 1,500,000 times, as shown in the last column of Table 2. The time to fixation increases when r is small, but shows no clear pattern when r is large; the effect is clearly minor.

Selection

The results with selection are presented in Tables 3 through 5. Our discussion focuses on the case of unidirectional mutation (Tables 3 and 4). Results with bidirectional mutation are completely analogous, as illustrated in Table 5 with multiplicative selection. Irrespective of the selection regime, both first appearance and first fixation times are longer with bidirectional mutation.

The time to first appearance declines with increasing recombination rates, showing the same qualitative behavior both with and without selection. This result holds for all population sizes, mutation rates and selection regimes tested (data not shown),

6

Table 2: Results in the absence of selection. Each combination of parameters was simulated 10,000 times (1,500,000 where noted). The time to first appearance of the double mutant (T_A) and the time to fixation (T_F) are given, with their standard errors in parentheses. Unidirectional mutation occurs at rate μ_F in a population of size N.

	Recombination Fraction	T_A	T_F	T_F 1,500,000 reps
$N = 100$	$r = 0.0$	10.49 (0.06)	251.50 (1.14)	252.17 (0.09)
$\mu_F = 0.01$	$r = 0.1$	9.54 (0.05)	251.30 (1.15)	252.93 (0.09)
	$r = 0.2$	8.99 (0.05)	253.22 (1.14)	253.09 (0.09)
	$r = 0.3$	8.68 (0.05)	252.60 (1.13)	253.00 (0.09)
	$r = 0.4$	8.35 (0.04)	252.66 (1.15)	253.06 (0.09)
	$r = 0.5$	8.18 (0.04)	254.55 (1.15)	253.09 (0.09)
$N = 1000$	$r = 0.0$	3.34 (0.02)	434.56 (1.24)	
$\mu_F = 0.01$	$r = 0.1$	3.27 (0.01)	435.18 (1.23)	
	$r = 0.2$	3.20 (0.01)	434.28 (1.22)	
	$r = 0.3$	3.12 (0.01)	433.56 (1.22)	
	$r = 0.4$	3.11 (0.01)	434.57 (1.23)	
	$r = 0.5$	3.03 (0.01)	432.69 (1.22)	
$N = 1000$	$r = 0.0$	30.63 (0.18)	2545.62 (11.58)	
$\mu_F = 0.001$	$r = 0.1$	24.56 (0.13)	2549.62 (11.41)	
	$r = 0.2$	21.73 (0.11)	2553.13 (11.61)	
	$r = 0.3$	20.10 (0.10)	2548.06 (11.38)	
	$r = 0.4$	18.92 (0.10)	2550.98 (11.61)	
	$r = 0.5$	18.14 (0.09)	2573.27 (11.75)	

although the effect is sometimes quite small. We have also observed that the time to first appearance does depend on the fitness of the double mutant. This rather troubling fact occurs because, in this model, selection acts deterministically on the *juvenile* stage of the life cycle but we census at the *adult* stage and then determine whether a double mutant is present. Were we to assume that selection occurred only among the individuals sampled to make the next generation, and to census the adults before selection, then clearly σ_1 would play no role in the first appearance of the double mutant.

The time to fixation behaves in a more complicated manner as a function of recombination. With submultiplicative selection and multiplicative selection, the time to fixation always decreases with increasing recombination. These results confirm the simulation results of Hill and Robertson (1966) and Felsenstein (1974). By continuity,

Table 3: Times to fixation with selection and unidirectional mutation. The population size is N and the forward mutation rate is μ_F. Each combination of parameters was simulated 10,000 times. The standard errors for T_F are given in parentheses.

		Time to fixation (T_F)		
	Recombination	$N = 100$	$N = 1000$	$N = 1000$
	Fraction	$\mu_F = 0.01$	$\mu_F = 0.01$	$\mu_F = 0.001$
	$r = 0.0$	22.43 (0.04)	26.71 (0.03)	36.33 (0.04)
$\sigma_2 = 2$	$r = 0.1$	20.96 (0.04)	25.66 (0.03)	31.08 (0.04)
$\sigma_1 = 3$	$r = 0.2$	20.29 (0.04)	25.05 (0.03)	30.01 (0.03)
$(\gamma = -1)$	$r = 0.3$	20.00 (0.04)	24.73 (0.03)	29.45 (0.03)
Submultiplicative	$r = 0.4$	19.80 (0.04)	24.46 (0.03)	29.04 (0.03)
	$r = 0.5$	19.58 (0.04)	24.25 (0.03)	28.93 (0.03)
	$r = 0.0$	15.19 (0.03)	17.32 (0.02)	24.19 (0.03)
$\sigma_2 = 2$	$r = 0.1$	14.66 (0.02)	17.26 (0.02)	21.76 (0.02)
$\sigma_1 = 4$	$r = 0.2$	14.44 (0.02)	17.21 (0.02)	21.35 (0.02)
$(\gamma = 0)$	$r = 0.3$	14.38 (0.02)	17.15 (0.02)	21.15 (0.02)
Multiplicative	$r = 0.4$	14.35 (0.02)	17.16 (0.02)	21.10 (0.02)
	$r = 0.5$	14.33 (0.02)	17.14 (0.02)	21.04 (0.02)
	$r = 0.0$	12.35 (0.02)	13.78 (0.02)	19.55 (0.02)
$\sigma_2 = 2$	$r = 0.1$	12.10 (0.02)	13.92 (0.02)	18.06 (0.02)
$\sigma_1 = 5$	$r = 0.2$	12.04 (0.02)	14.02 (0.02)	17.87 (0.02)
$(\gamma = 1)$	$r = 0.3$	12.06 (0.02)	14.12 (0.01)	17.81 (0.02)
Supermultiplicative	$r = 0.4$	12.10 (0.02)	14.17 (0.01)	17.84 (0.02)
	$r = 0.5$	12.13 (0.02)	14.22 (0.01)	17.91 (0.02)
	$r = 0.0$	8.10 (0.01)	8.65 (0.01)	12.83 (0.02)
$\sigma_2 = 2$	$r = 0.1$	8.14 (0.01)	8.89 (0.01)	12.38 (0.02)
$\sigma_1 = 10$	$r = 0.2$	8.18 (0.01)	9.09 (0.01)	12.38 (0.01)
$(\gamma = 6)$	$r = 0.3$	8.27 (0.01)	9.28 (0.01)	12.44 (0.01)
Supermultiplicative	$r = 0.4$	8.32 (0.01)	9.43 (0.01)	12.54 (0.01)
	$r = 0.5$	8.44 (0.01)	9.56 (0.01)	12.70 (0.01)
	$r = 0.0$	3.87 (0.01)	4.15 (0.01)	6.25 (0.01)
$\sigma_2 = 2$	$r = 0.1$	3.94 (0.01)	4.22 (0.01)	6.31 (0.01)
$\sigma_1 = 100$	$r = 0.2$	4.01 (0.01)	4.31 (0.01)	6.39 (0.01)
$(\gamma = 96)$	$r = 0.3$	4.08 (0.01)	4.40 (0.01)	6.47 (0.01)
Supermultiplicative	$r = 0.4$	4.16 (0.01)	4.50 (0.01)	6.52 (0.01)
	$r = 0.5$	4.22 (0.01)	4.61 (0.01)	6.56 (0.01)

Table 4: Times to first appearance with multiplicative selection and unidirectional mutation. The population size is N and the forward mutation rate is μ_F. Each combination of parameters was simulated 10,000 times. The standard errors for T_A are given in parentheses.

		Time to first appearance (T_A)		
	Recombination	$N = 100$	$N = 1000$	$N = 1000$
	Fraction	$\mu_F = 0.01$	$\mu_F = 0.01$	$\mu_F = 0.001$
	$r = 0.0$	3.73 (0.01)	1.78 (0.01)	6.37 (0.02)
$\sigma_2 = 2$	$r = 0.1$	3.50 (0.01)	1.77 (0.01)	5.41 (0.01)
$\sigma_1 = 4$	$r = 0.2$	3.38 (0.01)	1.77 (0.01)	5.12 (0.01)
($\gamma = 0$)	$r = 0.3$	3.31 (0.01)	1.75 (0.01)	4.91 (0.01)
Multiplicative	$r = 0.4$	3.23 (0.01)	1.75 (0.01)	4.80 (0.01)
	$r = 0.5$	3.19 (0.01)	1.73 (0.01)	4.69 (0.01)

we expect that recombination will continue to shorten the time necessary for populations to reach fixation with sufficiently weak supermultiplicative selection. With strong supermultiplicative selection, however, our simulations show that recombination actually increases the time to fixation. We interpret these results in the following manner. There is a tendency in finite populations to generate more negative linkage disequilibrium than expected on the basis of infinite population results. Whereas linkage disequilibrium becomes negative only for submultiplicative selection regimes in an infinite population, it becomes negative in finite populations whenever selection is submultiplicative, multiplicative, or weakly supermultiplicative. In fact, it can be shown analytically (see Hill and Robertson, 1966) that negative disequilibrium develops in a finite population with multiplicative selection. With sufficiently strong supermultiplicative selection, the positive disequilibrium created by selection is large enough to be maintained, despite the tendency for sampling with selection to make the disequilibrium more negative; in such cases, recombination delays fixation. That is, there exists some critical region of positive epistasis (an interval, $\{\gamma_c, \gamma_c'\}$, say) above which recombination increases the time to fixation, and below which recombination decreases the time to fixation. The simulations indicate that for epistasis within this interval the time to fixation might exhibit non-monotonic behavior with respect to the recombination rate. For instance, epistasis equal to $\gamma = 1$ appears to be below γ_c for populations of size $N = 1000$ with $\mu_F = 0.01$, but either above or within the critical region for populations of size $N = 100$ with $\mu_F = 0.01$ or of size $N = 1000$ with $\mu_F = 0.001$ We speculate that this critical region shifts toward zero as the population size increases (while the mutation rate is held constant), becoming zero at the limit of an infinitely large population.

9

Table 5: Results in the presence of multiplicative selection and bidirectional mutation. The population size is $N = 100$ and the forward and backward mutation rates equal $\mu_F = \mu_B = 0.01$. Each combination of parameters was simulated 10,000 times. The time to first appearance of the double mutant (T_A) and the time to fixation (T_F) are given, with their standard errors in parentheses.

	Recombination Fraction	T_A	T_F
	$r = 0.0$	3.78 (0.01)	22.12 (0.08)
$\sigma_2 = 2$	$r = 0.1$	3.54 (0.01)	21.68 (0.08)
$\sigma_1 = 4$	$r = 0.2$	3.39 (0.01)	21.39 (0.08)
$(\gamma = 0)$	$r = 0.3$	3.32 (0.01)	21.33 (0.08)
Multiplicative	$r = 0.4$	3.23 (0.01)	21.22 (0.08)
	$r = 0.5$	3.20 (0.01)	21.17 (0.08)

Time until Appearance of First Double Mutant.

In this section, we outline a proof that the time to first appearance of the double mutant must decline with increasing recombination.

The first appearance within the adult population of the double mutant (type ab) will occur in some generation (n, say) if in all $(n-1)$ previous generations, the ab type is not present, so that its frequency, x_1, equals zero. It follows that in any generation before n, the linkage disequilibrium in the population must equal zero or be negative $(D = x_1 x_4 - x_2 x_3 = -x_2 x_3 \leq 0)$.

In proceeding from generation k $(k < n)$ to generation $k+1$ according to the Wright-Fisher model, the probability that there is at least one ab type in the $k+1^{th}$ generation is $(P_{\{ab\}}^{k+1})$:

$$1 - (1 - x_1^{(k+1)})^N.$$

This is an increasing function of $x_1^{(k+1)}$, which is computed using equation (1) with the stipulation that $x_1^{(k)} = 0$:

$$x_1^{(k+1)} = \frac{\sigma_1 \{\mu_F (1 - \mu_B) x_2^{(k)} + \mu_F (1 - \mu_B) x_3^{(k)} + \mu_F^2 x_4^{(k)} - r D^{(k)} (1 - \mu_F - \mu_B)^2\}}{W},$$

where W is the sum of the right hand sides of equation (1) with $x_1^{(k)} = 0$. The derivative of $P_{\{ab\}}^{k+1}$ with respect to the recombination rate, r, has the same sign as the

derivative of $x_1^{(k+1)}$ with respect to r, which in turn has the same sign as:

$$-D\sigma_1\{(x_2^{(k)} + x_3^{(k)})(\sigma_2(1 + \mu_F - \mu_B) - (\mu_F - \mu_B)) + x_4^{(k)}(1 - 2\mu_F + 2\mu_F\sigma_2)\}. \quad (2)$$

For reasonable values of the parameters (fitnesses greater than the mutation rates), the derivative of $P_{\{ab\}}^{k+1}$ with respect to r simply has the opposite sign to the disequilibrium in the kth generation.

The probability of observing ab, as a function of the frequencies in the $k^{(th)}$ generation, thus has the following properties:

$$\begin{aligned}
\frac{\partial(P_{\{ab\}}^{k+1}|D^{(k)} = 0)}{\partial r} &= 0 \\
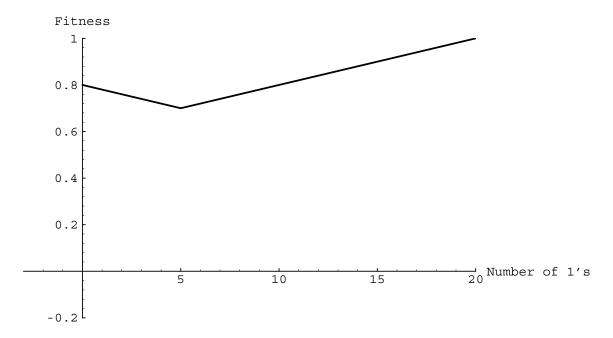\frac{\partial(P_{\{ab\}}^{k+1}|D^{(k)} < 0)}{\partial r} &> 0 \\
\frac{\partial(P_{\{ab\}}^{k+1}|D^{(k)} > 0)}{\partial r} &< 0
\end{aligned} \quad (3)$$

As mentioned above, the disequilibrium is either negative or zero in all generations before the first appearance of the double mutant. It follows that before the ab type has first appeared, the probability that there is a double mutant in the next generation is either higher with recombination or independent of recombination. Thus we expect recombination to decrease the time to first appearance of the double mutant in any finite population and in the limit, with an infinitely large population, recombination will have no influence on the appearance of the first double mutant (which will appear within one generation). This argument does not depend on the selection scheme or the population size.

In short, finite populations with higher rates of recombination will take less time on average to produce the first double mutant, *whether such double mutants are favorable or deleterious.* This result, however, depends critically on the odd conditioning involved in the measure of first appearance. In a finite population with mutation, the expected disequilibrium, conditional on the fact that the first double mutant has not appeared, must always be negative. Since the observed populations are restricted to having non-positive disequilibria, the time to first appearance must decrease with recombination. Thus the time to first appearance is not a very good diagnostic tool for determining whether recombination is advantageous, because it can only give one answer.

Multi-locus Simulations

Figure 1: Fitness regime for the twenty-locus simulations. The above function gives the fitness of an individual according to the total number of 1 alleles at all loci.

The above results can be extended to many loci using the simulation methods known as genetic algorithms. In such simulations, the fitness regimes may be quite general. As an example, we consider a twenty locus model with two alleles (0/1) at each locus. The fitness of an individual depends on the number of 1's in its genome as illustrated in Figure 1. With this fitness scheme and starting from a population fixed at all loci on the 0 allele , we find that both the time to first appearance of a string of all 1's and the time to fixation of this string *increases* with recombination (Table 6). In this multi-locus case, measuring the time to first appearance does not restrict the pairwise disequilibria to a non-positive value so that recombination can slow the appearance of the most fit genotype. Different results are obtained, however, when the initial proportion of 1's is large *even though the initial linkage disequilibrium is zero*. When there are an equal proportion of 0 and 1 alleles, recombination hastens both the first appearance and the fixation of the string of all 1's. This indicates that initial allele frequencies, in the absence of linkage disequilibrium, are also critical in determining whether or not recombination is advantageous. We are in the process of extending these multi-locus simulations to produce a more complete picture of the role of selection and of initial population composition on the advantages and disadvantages of recombination.

Table 6: Twenty-locus results with fitness specified in Figure 1. In the simulations, the population size was $N = 100$, the number of recombination events in the entire population with 20 loci was set at R, and the number of forward mutations in the entire population of 20 loci was set at $m_F = 2$. The initial population composition was 100 strings of 20 loci with the given percentage of $1's$ randomly placed (expected linkage disequilibrium of zero). Each combination of parameters was simulated 500 times. The time to first appearance of the string with all $1's$ (T_A) and the time to fixation of this string (T_F) are given, with their standard errors in parentheses.

Percentage of 1's	Recombination Number	T_A	T_F
0%	$R = 0$	2015.00 (27.15)	2425.46 (29.58)
	$R = 1$	2257.10 (36.42)	2650.81 (35.99)
	$R = 10$	3479.76 (73.58)	3865.62 (76.42)
	$R = 30$	3916.70 (88.44)	4072.70 (73.61)
10%	$R = 0$	1776.06 (29.25)	2241.05 (31.22)
	$R = 1$	2043.18 (34.23)	2501.68 (37.04)
	$R = 10$	3203.26 (66.28)	3482.26 (65.23)
	$R = 30$	3834.41 (90.66)	4164.34 (85.66)
25%	$R = 0$	756.37 (15.94)	1236.67 (24.16)
	$R = 1$	776.64 (23.22)	1207.53 (27.91)
	$R = 10$	880.93 (49.53)	1428.63 (56.26)
	$R = 30$	895.49 (57.46)	1355.66 (58.73)
50%	$R = 0$	451.95 (8.15)	881.89 (15.41)
	$R = 1$	376.50 (7.28)	793.85 (14.36)
	$R = 10$	223.77 (5.89)	662.56 (14.98)
	$R = 30$	155.18 (5.25)	592.22 (14.57)

CONCLUSIONS

The advantages or disadvantages of recombination depend critically on the kind of model studied. While we have shown that recombination will always shorten the time to first appearance of a double mutant in the two-locus Wright-Fisher model, it may or may not shorten the time to fixation of this chromosome. In fact, the time to fixation is particularly sensitive to the assumptions made. Table 7 summarizes qualitatively the effect of recombination on the time to fixation of a favorable double mutant in a population that is initially fixed on the non-mutant chromosome. From our discussion above, it is clear that different results can be obtained if the population is started in linkage disequilibrium. Roughly, if the population is initially in negative disequilibrium, the advantage of recombination is increased (or its disadvantage is

Table 7: Summary of the conclusions. Populations are initially in linkage equilibrium with no mutant alleles. In the absence of selection in finite populations, the increase in T_F with r is *very* slight and we assume that selection is sufficiently strong to counter this effect.

Selection	Finite Population T_F	Infinite Population T_F
None	↑ with r	Independent of r
Submultiplicative	↓ with r	↓ with r
Multiplicative	↓ with r	Independent of r
Supermultiplicative ($\gamma < \gamma_c$)	↓ with r	↑ with r
Supermultiplicative ($\gamma > \gamma_c'$)	↑ with r	↑ with r

decreased). Conversely, if the population commences with positive disequilibrium, the advantage of recombination is decreased (or its disadvantage is increased).

Our claim is that the behavior of linkage disequilibrium is particularly useful in determining when recombination might be advantageous. This runs counter to the views of Hill and Robertson (1966), who said that disequilibrium was not particularly "illuminating." Remember, however, that an understanding of the development of disequilibria helped us to predict the existence of a critical region ($\{\gamma_c, \gamma_c'\}$) in the amount of positive epistasis above which T_F would increase with recombination. Concentration on the effective population size alone, as suggested by Hill and Robertson (1966), would not have led to such a prediction. It is also more difficult to predict the role of initial population composition using effective population size as a guide. We suggest that both the effective population size and the disequilibrium be used to interpret models that attempt to explain the advantage of recombination.

Results from simulations similar to those used in the implementation of genetic algorithms promise to extend our understanding of the role of recombination to more loci and more general fitness schemes. Using one possible fitness scheme (Figure 1), we observed that both the time to first appearance and the time to fixation of the most fit genotype can increase or decrease with recombination depending on initial allele frequencies. More generally, Bergman and Feldman (1990, 1992) have shown that the trajectory of a recombination modifying allele in a multilocus system depends strongly on the shape of the fitness regime. Indeed, there is an increasing body of evidence (e.g. Forrest and Mitchell, 1991) that the more jagged is the fitness as a function of the genotypic value, the less advantage there is to recombination in accelerating the appearance of the fittest type. This aspect of the "recombination

question" requires further study.

ACKNOWLEDGEMENTS

BIBLIOGRAPHY

BELEW, R. K. AND BOOKER, L. B. 1991. Proceedings of the Fourth International Conference on Genetic Algorithms. Morgan Kaufmann, San Mateo, CA.

BERGMAN, A. AND FELDMAN, M. 1990. More on selction for and against recombination. Theor. Pop. Biol. 38: 68–92.

BERGMAN, A. AND FELDMAN, M. 1992. Recombination dynamics and the fitness landscape. Physica D 56: 57–67.

BODMER, W. F. 1970. The evolutionary significance recombination in prokayotes. Symp. Soc. Gen. Microbiol. 20: 279–294.

CROW, J. F. AND KIMURA, M. 1965. Evolution in sexual and asexual populations. Am. Nat. 99: 439–450.

ESHEL, I. AND FELDMAN, M. W. 1970. On the evolutionary effect of recombination. Theor. Pop. Biol. 1: 88–100.

FELSENSTEIN, J. 1974. The evolutionary advantage of recombination. Genetics 78: 737–756.

FELSENSTEIN, J. 1988. Sex and the evolution of recombination. *In* R. E. Michod and B. R. Levin (Eds.), The Evolution of Sex: An Examination of Current Ideas (pp. 74–86). Sinauer, Sunderland, MA.

FELSENSTEIN, J. AND YOKOYAMA, S. 1976. The evolutionary advantage of recombination. II. Individual selection for recombination. Genetics 83: 845–859.

FORREST, S. AND MITCHELL, M. 1991. The performance of genetic algorithms on Walsh polynomials: Some anomalous results and their explanation. *In* R. K. Belew and L. B. Booker (Eds.), Proceedings of the Fourth International Conference on Genetic Algorithms (pp. 182–189). Morgan Kaufmann, San Mateo, CA.

GOLDBERG, D. E. 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA.

HAIGH, J. 1978. The accumulation of deleterious genes in a population: Muller's ratchet. Theor. Pop. Biol. 14: 251–267.

HILL, W. G. AND ROBERTSON, A. 1966. The effect of linkage on the limits to artificial selection. Genet. Res. Camb. 8: 269–294.

HOLLAND, J. H. 1992. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications in Biology, Control, and Artificial Intelligence. MIT Press, Cambridge, MA.

KARLIN, S. 1973. Sex and infinity: a mathematical analysis of the advantages and disadvantages of genetic recombination. *In* M. S. B. and R. W. Hiorns (Ed.), The Mathematical Theory of the Dynamics of Biological Populations (pp. 155–194). Academic Press, New York.

MAYNARD-SMITH, J. 1968. Evolution in sexual and asexual populations. Am. Nat. 102: 469–473.

SCHAFFER, J. D. AND ESHELMAN, L. J. 1991. On crossover as an evolutionary viable strategy. *In* R. K. Belew and L. B. Booker (Eds.), Proceedings of the Fourth International Conference on Genetic Algorithms (pp. 61–68). Morgan Kaufmann, San Mateo, CA.

SUMIDA, B. H., HOUSTON, A. I., MCNAMARA, J. M., AND HAMILTON, W. D. 1990. Genetic algorithms and evolution. J. Theor. Biol. 147: 59–84.