SELECTION FOR RECOMBINATION IN SMALL POPULATIONS

SARAH P. OTTO[1,2] AND NICK H. BARTON[3,4]

[1]*Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada*
[2]*E-mail: otto@zoology.ubc.ca*
[3]*Institute of Cell, Animal and Population Biology, West Mains Road, Edinburgh EH9 3JT, United Kingdom*
[4]*E-mail: n.barton@ed.ac.uk*

Abstract.—The reasons that sex and recombination are so widespread remain elusive. One popular hypothesis is that sex and recombination promote adaptation to a changing environment. The strongest evidence that increased recombination may evolve because recombination promotes adaptation comes from artificially selected populations. Recombination rates have been found to increase as a correlated response to selection on traits unrelated to recombination in several artificial selection experiments and in a comparison of domesticated and nondomesticated mammals. There are, however, several alternative explanations for the increase in recombination in such populations, including two different evolutionary explanations. The first is that the form of selection is epistatic, generating linkage disequilibria among selected loci, which can indirectly favor modifier alleles that increase recombination. The second is that random genetic drift in selected populations tends to generate disequilibria such that beneficial alleles are often found in different individuals; modifier alleles that increase recombination can bring together such favorable alleles and thus may be found in individuals with greater fitness. In this paper, we compare the evolutionary forces acting on recombination in finite populations subject to strong selection. To our surprise, we found that drift accounted for the majority of selection for increased recombination observed in simulations of small to moderately large populations, suggesting that, unless selected populations are large, epistasis plays a secondary role in the evolution of recombination.

Key words.—Drift, epistasis, Fisher-Muller theory, modifier, recombination.

Producing offspring through sexual recombination is a risky endeavor. An organism must find and attract a mate, while simultaneously risking the attraction of predators and parasites. An organism must invest in male as well as female structures or offspring, even when such a division of resources is inefficient. Perhaps most puzzling of all, an organism must undergo meiosis and syngamy, during which the genotype of the parent, who has survived to reproduce under current environmental conditions, is recombined and mixed with the genotype of another, with no guarantee that the resulting offspring will carry a successful combination of alleles. The ubiquity of sex and recombination suggests, however, that the long-term benefits of genomic mixing do outweigh the risks. Numerous theoretical studies have clarified these benefits and specified the conditions under which sexual reproduction and higher rates of recombination might evolve (see reviews by Michod and Levin 1988; Korol et al. 1994; Barton and Charlesworth 1998; Otto and Michalakis 1998; West et al. 1999).

One of the most favorable circumstances for the evolution of recombination arises when a population is subject to directional selection, as might occur in a novel environment. Under such circumstances, sex and recombination can increase the amount of genetic variation and improve the response to selection of a population (see review by Burt 2000). Theoretical analyses have demonstrated, however, that directional selection does not always favor the evolution of higher rates of recombination (Maynard Smith 1980, 1988; Charlesworth 1993a; Barton 1995a). Selection for increased recombination is found in deterministic models (i.e., assuming very large populations with numerous copies of each allele such that random genetic drift is negligible) only when there are fitness interactions among loci of a particular sort. Specifically, the fitness of individuals carrying two beneficial alleles together must be less than the product of the fitnesses of individuals carrying each alone (negative epistasis) for recombination to hasten the spread of the beneficial alleles. Even then, increased rates of recombination will only evolve if epistasis is sufficiently weak and negative because only in this case are the long-term benefits of recombination greater than the immediate costs of producing recombinant offspring (Barton 1995a).

It is worth briefly reviewing the mechanisms underlying this result (see also Fig. 1). Negative epistasis generates negative linkage disequilibria among favorable alleles, such that beneficial alleles tend to be found on genetic backgrounds that are less fit than average (Eshel and Feldman 1970). With negative disequilibria among favorable alleles, recombination mixes up the genetic backgrounds and produces offspring with a broader array of fitnesses. Some recombinants will carry a larger number of beneficial alleles, whereas other recombinants will be burdened with an excess of deleterious alleles. Because epistasis is negative, however, the amount by which the former offspring have a higher fitness tends to be less than the amount by which the latter offspring have a lower fitness when compared to the fitness of nonrecombinant offspring (Charlesworth and Barton 1996). Nevertheless, those recombinants carrying many beneficial mutations will, in the long-term, be more likely to reproduce, thereby propagating alleles that promote recombination. Consequently, modifier alleles that increase recombination only spread when epistasis is negative and weak enough that the immediate fitness consequences of recombination are mild compared to the long-term advantage of bringing together beneficial alleles onto the same genetic background.

There is indeed good evidence that novel environments do favor the evolution of recombination. Table 1 presents a re-
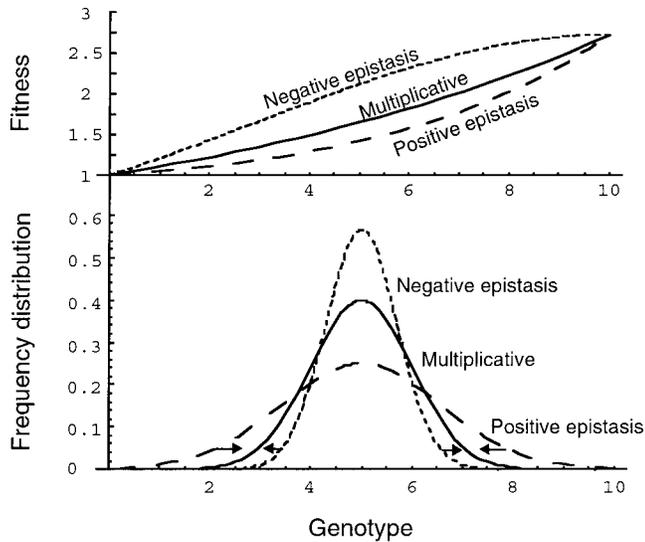
FIG. 1. The effects of epistasis. In the top panel, fitness is plotted as a function of genotype (x) using the fitness function $e^{sx+cs^2x^2}$. When $c = 0$, fitness is multiplicative (solid curves); when $c < 0$, there is negative epistasis ($c = -1/4$ is illustrated by short dashed curves); and when $c > 0$, there is positive epistasis ($c = 4$ is illustrated by long dashed curves). In the bottom panel, the frequency distribution of genotypes, x, is shown after 50 generations of selection for an asexual population whose initial frequency distribution is Gaussian, with a mean of zero and a variance of one. The strength of selection (s) is standardized for each value of c such that the mean genotype after 50 generations of selection is five. Mutation is ignored. With negative epistasis, intermediate genotypes have a relatively high fitness compared to extreme genotypes; consequently, the frequency distribution after selection exhibits less variance than when there is multiplicative epistasis. A new modifier allele that introduces recombination into such a population will increase the genotypic variance (see arrows), producing more of the fittest genotypes and improving the ability of the population to respond to further selection. This provides a long-term benefit to the modifier. With positive epistasis, selection itself increases the genotypic variance. In this case, recombination would reduce the variance (see arrows), causing a long-term disadvantage to modifiers that increase recombination. In both cases illustrated, however, the genotypes that become common in response to selection are fitter, on average, than the genotypes produced by recombination. Over the short-term, this phenomenon, known as the recombination load, selects against modifiers that increase the rate of recombination. Overall, selection on a modifier of recombination reflects a balance between these long-term and short-term effects (Barton 1995a).

view of studies in which recombination has been measured during or following a period of artificial selection. The correlated response in recombination to artificial selection varies in both magnitude and direction, but the majority of cases exhibit increased recombination rates. Furthermore, the magnitude of change was often substantial, with local increases in recombination of over 30% observed in three independent studies in *Drosophila melanogaster* (Flexon and Rodell 1982; Zhuchenko et al. 1985; Korol and Iliadi 1994).

The fact that increased recombination is observed in such selection experiments has been taken as evidence that the right form of epistasis—weak and negative—is prevalent. Yet, there is another possible explanation. The fittest genotypes may be underrepresented or may simply not exist within small populations under directional selection. Recombination

could generate these genotypes by combining beneficial alleles that are carried within different individuals. This advantage is the basis of several classical arguments for the origin and persistence of sex and recombination (Morgan 1913; Fisher 1930; Muller 1932; Haldane 1990), and it relies on the stochastic nature of evolution in small populations. In infinitely large populations, every genotype already exists; thus, as long as there is no epistasis and no initial linkage disequilibrium, genes that modify recombination rates experience no selection (Maynard Smith 1968). Intuitively, one might argue that random genetic drift should be unbiased, increasing or decreasing the frequency of beneficial gene combinations to an equal degree. Although this is true, the joint action of drift and selection is biased and tends to generate negative linkage disequilibria among favorable alleles (Hill and Robertson 1966; Felsenstein 1974). When, by chance, beneficial alleles are found on the same chromosome (positive linkage disequilibria), selection becomes more efficient, leading to the rapid rise in frequency of the favored chromosome and to the rapid dissipation of linkage disequilibria. When beneficial alleles occur by chance on poor genetic backgrounds (negative linkage disequilibria), however, selection slows down, essentially freezing the population in a state of negative disequilibria. Recently, we have quantified the amount of disequilibrium that accumulates over time due to the combined action of drift and recombination (N. H. Barton and S. P. Otto, unpubl. ms.) and have determined how selection and disequilibria affect the fixation probability of beneficial mutations (Barton 1995b). Furthermore, we have shown that when linkage disequilibria are randomly generated (either by the random appearance of mutations on particular backgrounds or by drift in small populations), modifier alleles that increase recombination are positively selected, even if there is no epistasis and no initial linkage disequilibrium (Otto and Barton 1997; N. H. Barton and S. P. Otto, unpubl. ms.).

Unfortunately, the joint effects of drift and selection are difficult to model, and analytical progress has been made only by making several simplifying assumptions, including that selection is weak, populations are relatively large, and modifier genes have only slight effects on recombination. Under these assumptions, a gene that modifies recombination by a proportion, δr, experiences a selection coefficient of approximately $\delta r\, V_1\, V_2/(2Nr^4)$ at quasi-linkage equilibrium (QLE) in a diploid population of size N, given sufficiently loose linkage (r) between the modifier and a pair of selected loci with additive genetic variances in fitness, V_1 and V_2, and no epistasis (N. H. Barton and S. P. Otto, unpubl. ms.). Yet, the assumptions of the model prevent us from knowing the role that drift may play in the evolution of recombination observed in cases, such as those described in Table 1, where selection is strong and population sizes are small. It seems plausible that selection on recombination will be greatest in such cases, even if they are less mathematically tractable and thus relatively neglected.

Here we use simulation to examine the forces acting on recombination rates in small populations subject to strong selection. We compare results with and without epistasis in populations of varying size to determine the extent to which the evolution of recombination is driven by epistasis or by

TABLE 1. Effect of artificial selection on recombination. Degree of recombination is estimated from recombinants among visible markers (A–F), chiasma counts (F–H0), or sporulation rates (I). See Korol et al. (1994) and Korol and Iliadi (1994) for further review.

Study[1]	Organism	Selection regime	N[2]	Time[3]	Effect on recombination[4]		
A	*Drosophila melanogaster*	resistance to DDT	? (<100)	22–38	−0.7% (on X)	+18.9% (on 2)*	+67.4% (on 3)*
B	*D. melanogaster*	daily temperature fluctuations	<6000	<80	+19.6% (on X)*	+25.1% (on 2)*	+32.5% (on 3)*
C	*D. melanogaster*	positive geotaxis	30	50	+51.8% (on X)*	+33.9% (on 2)*	+21.3% (on 3)*
C	*D. melanogaster*	negative geotaxis	30	50	+71.7% (on X)*	−1.5% (on 2)	+25.7% (on 3)*
D	*D. melanogaster*	wing length in both sexes	?	50	−17.2% (on 2)*		
D	*D. melanogaster*	wing length in females	?	50	+6.1% (on 2)*		
E	*D. melanogaster*	high motor activity	?	?	−6.8% (on 2)		
E	*D. melanogaster*	low motor activity	?	?	−17.1% (on 2)		
F	mice	large body weight	160	59	+2.6%*		
F	mice	large litter size	160	74	−3.5%*		
G	seven mammal species	domestication complex	?	?	+100%[5]		
H	mustard	early flowering time	?	5	−1.3% (year 1)	+2.5% (year 2)	[self-compatible lines]
					+13.4% (year 1)*	−4.4% (year 2)	[self-incompatible lines]
H	mustard	late flowering time	?	5	+1.0% (year 1)	+8.1% (year 2)	[self-compatible lines]
					+18.5% (year 1)	−0.1% (year 2)	[self-incompatible lines]
I	yeast	nutrient fluctuations	>5 × 10[5]	36	Control:	ns (1st half)	ns (2nd half)
					Weekly fluct:	− (1st half)	ns (2nd half)
					Mild fluct:	ns (1st half)	ns (2nd half)
					Weekly fluct:	ns (1st half)	ns (2nd half)
					Biweekly fluct:	+ (1st half)	− (2nd half)
					Daily fluct:	+ (1st half)	ns (2nd half)

[1] Studies: (A) Flexon and Rodell (1982); (B) Gorodetskii et al. (1990); results from additional experiments on adaptation to fluctuating temperature by this group are reviewed by Korol et al. (1994); (C) Korol and Iliadi (1994); (D) Palenzona et al. (1972); (E) Lobashev et al. (1973); (F) Gorlov et al. (1992); (G) Burt and Bell (1987); (H) Harinarayana and Murty (1971); *t*-test based on a comparison of a single control mean with replicate line means; (I) Wolf et al. (1987).

[2] Diploid population size of selected lines (?, not reported, followed by an educated guess); note that the effective population sizes, which should be used for comparison with the simulations, will generally be smaller.

[3] Duration of selection experiment measured in generations except for yeast, which is measured in weeks with ~40 asexual and one sexual generation per week.

[4] Change in recombination reported as percent change in recombination compared to controls (* significant at $P < 0.05$), except for yeast, whose results are described as a significant increase (+), decrease (−), or no significant change in recombination (ns) at an α-level of 0.05.

[5] Approximate average increase in the number of chiasmata in excess of one per chromosome.

drift. In particular, we ask whether disequilibrium generated by epistasis or by drift is more likely to account for the changes in recombination that have been observed experimentally. Finally, we suggest experiments to distinguish the relative contributions of these two sources of disequilibria.

METHODS

Two simulation programs were developed. The first was based on a three-locus model, with one locus modifying recombination (M) and two subject to selection (T_1 and T_2), in the order M T_1 T_2. A constant population of $2N$ chromosomes was simulated, corresponding to a diploid population of N individuals or a haploid population of $2N$ individuals. Each generation, a new population of chromosomes was randomly sampled with replacement from a multinomial distribution, where the chance of sampling each chromosome type was set to its expected frequency after selection followed by recombination using the recursions in Feldman (1972). Labeling deleterious and beneficial alleles as $-$ and $+$, respectively, selection was assumed to depend on the haplotype of a chromosome, with fitness equal to 1, $1 + s$, $1 + s$, $(1 + s)^2 + \epsilon$ for the $--$, $+-$, $-+$, and $++$ haplotypes, respectively. Epistasis, in this model, is measured by ϵ. Although this model of epistasis is natural for haploids, in diploids, it allows for interactions only among gene products encoded on the same chromosome and not on its homologue. Nevertheless, this simple model captures the essential manner by which epistasis generates selection on modifiers of recombination.

The second program allowed for an arbitrary number of loci by keeping track of chromosomes as strings of pluses and minuses. Each generation, chromosomes were generated by randomly choosing two chromosomes from the previous generation and allowing recombination to occur between adjacent pairs of loci without interference. The fitness of such chromosomes was assessed, and either multiplicative or truncation selection was applied. Many artificial selection experiments employ truncation selection by allowing individuals to reproduce only if their phenotype exceeds some threshold (e.g., the studies C, D, E, F, and H in Table 1). Truncation selection is also epistatic, and it can generate disequilibria that favor the evolution of increased recombination (Charlesworth 1993a; Kondrashov and Yampolsky 1996). To determine the relative importance of drift versus epistatic truncation selection, we compared simulations under multiplicative and truncation selection across a range of population sizes. With multiplicative selection, the fitness of a chromosome with j plus alleles at n selected loci equaled $(1 + s)^j/(1 + s)^n$; the chromosome would survive if a random number chosen uniformly between zero and one was less than its fitness, and it would die otherwise. This process was repeated until $2N$ chromosomes survived to comprise the next generation. To simulate truncation selection, $2N\tau$ chromosomes were examined each generation, from which the $2N$ chromosomes with the most pluses were allowed to survive. Truncation selection was made stronger by increasing τ and thus decreasing the proportion of the population allowed to reproduce ($1/\tau$). In addition to the selected loci, there was one modifier locus at the end of the chromosome.

In both programs, the initial population was formed by randomly setting p_0 of the selected alleles to $+$ and the remainder to $-$. Similarly, at the modifier loci, half of the alleles were randomly set to $+$. This procedure ensured that, on average, there was no linkage disequilibrium among loci at the beginning of the simulations. The initial average recombination rate between adjacent loci was set to $\bar{r}$. The exact recombination rate for an individual depended on the modifier genotype in the diploid phase, with recombination rates set to $\bar{r}(1 - \delta r)$, $\bar{r}$, and $\bar{r}(1 + \delta r)$ for $--$, $+-$, and $++$ individuals at the M locus, respectively. In the first program, only the recombination rate between the selected loci was modified, and the recombination rate between the modifier and the first selected locus was set arbitrarily to $\bar{r}(1 - \delta r)$. In the second program, the same recombination rate was used between any two adjacent loci (including modifier loci). Mutation was ignored throughout, based on the assumption that the majority of response to short-term selection depends on standing genetic variance rather than new mutations. The simulations were run for 50 generations, at which point the proportional change in recombination rate was recorded.

RESULTS

The results from simulations with three loci and strong selection ($s = 1$) are presented in Figure 2. With multiplicative selection (solid curves), the change in recombination is expected to equal zero for very small ($2N \to 1$) and very large ($2N \to \infty$) populations, but it becomes positive for populations of intermediate size. The peak response in recombination occurs in the range of $2N = 500$–1000 when beneficial alleles were initially rare ($p_0 = 0.01$) and in the range $2N = 50$–100 when beneficial alleles were initially more common ($p_0 = 0.1$). To test whether the results were sensitive to the assumption that the initial modifier frequency was 0.5, we ran additional simulations with an initial modifier frequency of 0.1 and 0.9. When the average recombination rate and the variance in recombination were set to the same initial values, similar proportional changes in recombination were observed. Surprisingly, adding negative epistasis made little difference to the extent of recombination evolution except in large populations ($2N > 10,000$ for $p_0 = 0.01$ and $2N > 100$ for $p_0 = 0.1$). Furthermore, even in large populations, epistasis generated strong selection for recombination only when the level of epistasis was at or near the best-case level of epistasis (Fig. 2, long dashed curves), that is, the level that maximizes the change in recombination rate in deterministic simulations ($2N = \infty$). Overall, the evolution of recombination in simulations with $\epsilon = -0.2$ (Fig. 2, short dashed curves) was similar to that observed with no epistasis. Additional support for the claim that selection for recombination is primarily driven by drift rather than epistasis in small populations comes from simulations with positive epistasis ($\epsilon = +1.0$; Fig. 2, thin solid curves). In an infinite population, positive epistasis generates disequilibria that select against recombination (Fig. 1). Indeed, in simulations with large populations, a decline in the average rate of recombination was observed. In small populations, however, increased rates of recombination were observed, at levels similar to simulations without epistasis, indicating that disequilibrium generated by
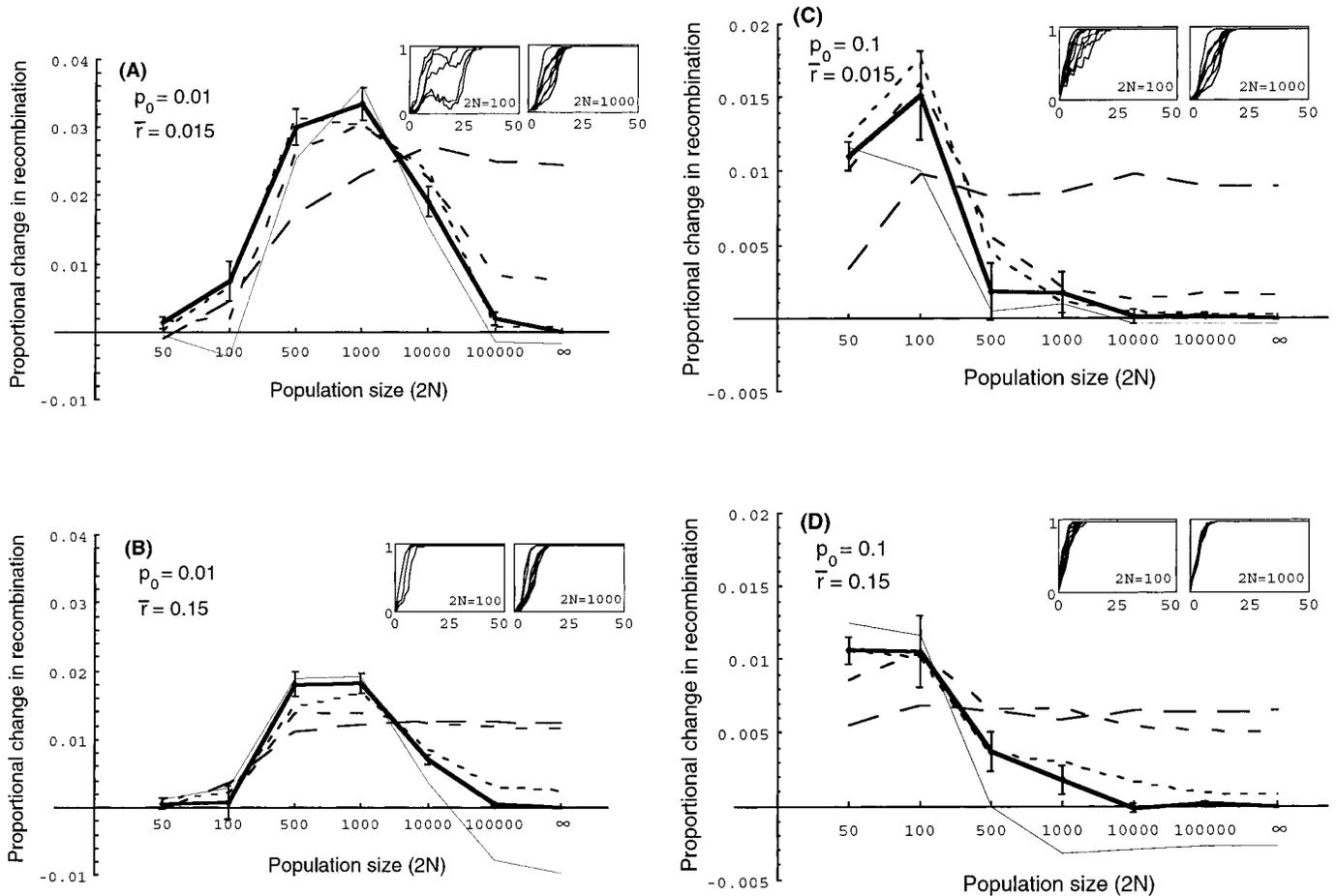
FIG. 2. The proportional change in recombination as a function of population size after 50 generations of selection in the three-locus model. The average frequency of the plus allele at the modifier locus was determined from 10^4 replicate runs (10^5 for $2N = 50$), and the proportional change in recombination over the 50 generations was calculated. In each case, strong selection with $s = 1$ was assumed, and δr was set to 1/3 such that $++$ individuals at the modifier locus underwent twice as much recombination as $--$ individuals. The curves illustrate results with varying degrees of negative epistasis: none (thick solid curves; ± 1 SE); weak ($\epsilon = -0.2$; short dashed curves); moderate ($\epsilon = -1.0$; medium dashed curves); and best case (long dashed curves), where the best-case level of epistasis was the one that caused the greatest increase in recombination in deterministic simulations ($2N = \infty$). For contrast, results from simulations with positive epistasis are also shown ($\epsilon = +1.0$; thin solid curves). Standard error bars are shown only for the case of no epistasis but are comparable in magnitude for the other cases. Inset figures show the change in allele frequency at one of the selected loci for $2N = 100$ and 1000, based on 10 replicates with multiplicative selection. (A) Low initial allele frequency ($p_0 = 0.01$) and tight linkage ($\bar{r} = 0.015$), giving a best-case $\epsilon = -1.69$. (B) Low initial allele frequency ($p_0 = 0.01$) and loose linkage ($\bar{r} = 0.15$), giving a best-case $\epsilon = -1.21$. (C) High initial allele frequency ($p_0 = 0.1$) and tight linkage ($\bar{r} = 0.015$), giving a best-case $\epsilon = -1.80$. (D) High initial allele frequency ($p_0 = 0.1$) and loose linkage ($\bar{r} = 0.15$), giving a best-case $\epsilon = -1.39$.

drift remains a powerful force selecting for recombination in small populations even with positive epistasis.

In fact, the simulations suggest that drift can cause a reversal in the roles of positive and negative epistasis. This reversal is most clearly seen in Figure 2B at intermediate population sizes: The observed increase in recombination is greatest with positive epistasis, followed by multiplicative selection, and finally by negative epistasis. This interesting result occurs because, in small populations, disequilibria may be generated more by drift than by epistasis. As described earlier, drift in the presence of selection tends to generate negative disequilibria, such that beneficial alleles tend to be found on genetic backgrounds with relatively low fitness. If drift is strong enough to generate negative disequilibria even in the face of positive epistasis, a modifier that increases

recombination gains both a short-term and a long-term advantage. As before, the long-term advantage arises when disequilibria are negative because recombination generates more of the extreme genotypes, thus increasing the variance in fitness upon which selection can act. The short-term advantage arises because, under positive epistasis, the extreme genotypes have a higher mean fitness than the intermediate genotypes. This short-term advantage decreases and eventually becomes negative as epistasis goes from being positive to negative. Thus, given that negative disequilibria are present, positive epistasis promotes the evolution of recombination, whereas negative epistasis hinders it (for similar results when migration generates negative disequilibria, see Lenormand and Otto 2000). For the simulations shown in Figure 2B with positive epistasis, the disequilibrium between
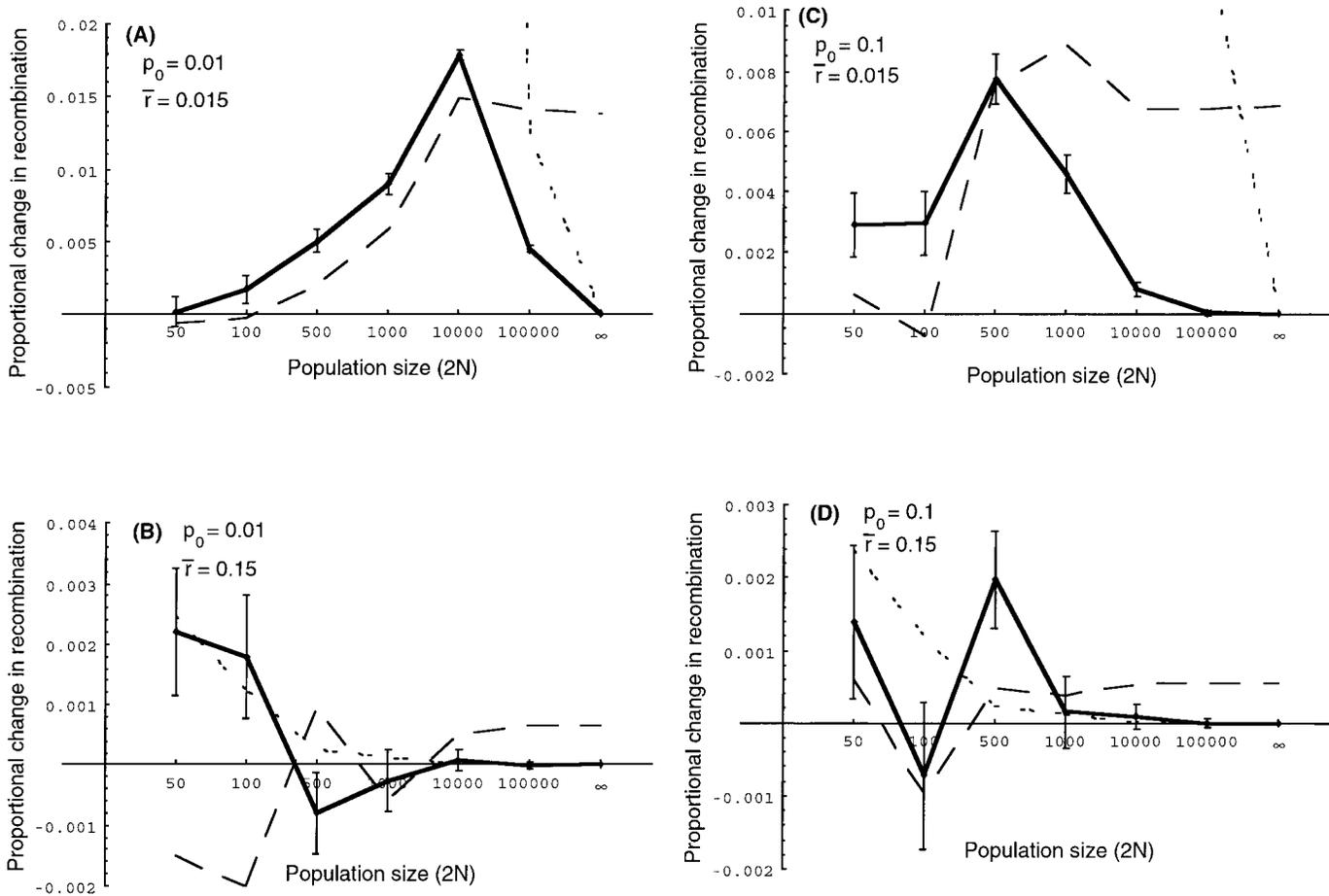
FIG. 3. The proportional change in recombination as a function of population size in the three-locus model with weaker selection. The simulations were identical to those shown in Figure 2, but with weaker selection ($s = 0.1$), a longer simulation time (250 generations), and 10^5 replicates per parameter set. Only multiplicative selection (thick solid curves; ± 1 SE) and the best-case level of epistasis (long dashed curves) are shown, alongside the quasi-linkage equilibrium prediction (double dashed curves) from N. H. Barton and S. P. Otto (unpubl. ms.). (A) Low initial allele frequency ($p_0 = 0.01$) and tight linkage ($\bar{r} = 0.015$), giving a best-case $\epsilon = -0.064$. (B) Low initial allele frequency ($p_0 = 0.01$) and loose linkage ($\bar{r} = 0.15$), giving a best-case $\epsilon = -0.033$. (C) High initial allele frequency ($p_0 = 0.1$) and tight linkage ($\bar{r} = 0.015$), giving a best-case $\epsilon = -0.075$. (D) High initial allele frequency ($p_0 = 0.1$) and loose linkage ($\bar{r} = 0.15$), giving a best-case $\epsilon = -0.033$.

the two selected loci is negative for $2N = 50$, 100, and 500; near zero, on average, for $2N = 1000$; and positive for all larger populations. These results suggest that, for these parameters, $2N \approx 1000$ forms the boundary between where drift and epistasis dominate the production of disequilibria.

One issue that makes comparisons among levels of epistasis difficult is that the additive genetic variance in fitness, $V_A(W)$, depends on the fitness regime and thus on the level of epistasis. At the beginning of the simulations, the plus alleles at the selected loci were relatively rare, and $V_A(W)$ was very similar for all levels of epistasis. When the plus alleles became more common, however, $V_A(W)$ was smaller in simulations with more negative levels of epistasis. Nevertheless, with $\epsilon = -0.2$ and $\epsilon = -1$, the total change in allele frequency at the selected loci (Δp) over the 50 generations was generally within 1% and always within 5% of the change in allele frequency observed with multiplicative selection. This was also true with loose linkage when ϵ was set to its best-case value, but with tight linkage the best-case level of epistasis was very strong, which slowed selection

down such that the beneficial alleles were, on average, only 92.6% as frequent after 50 generations.

Although the above results assumed very strong selection, increased recombination also evolves under weaker selection with drift. For comparison, we repeated the above simulations with $s = 0.1$ (Fig. 3), allowing more time (250 generations) for selection to act. The results with tight linkage were qualitatively similar. The percent change in recombination was lower but the same order of magnitude, although the peak effect occurred at larger population sizes (Figs 3A, 3C). With loose recombination, however, the percent change in recombination was much lower and was not significantly different from zero for most parameter sets (Figs 3B, 3D). Nevertheless, except in large populations, the presence of the best-case level of epistasis (Fig. 3, long dashed curves) did not promote the evolution of recombination any more than multiplicative selection.

These simulations allow a comparison with the QLE results obtained by us (unpubl. ms.). When recombination rates are lower than the selection coefficients, the QLE assumptions
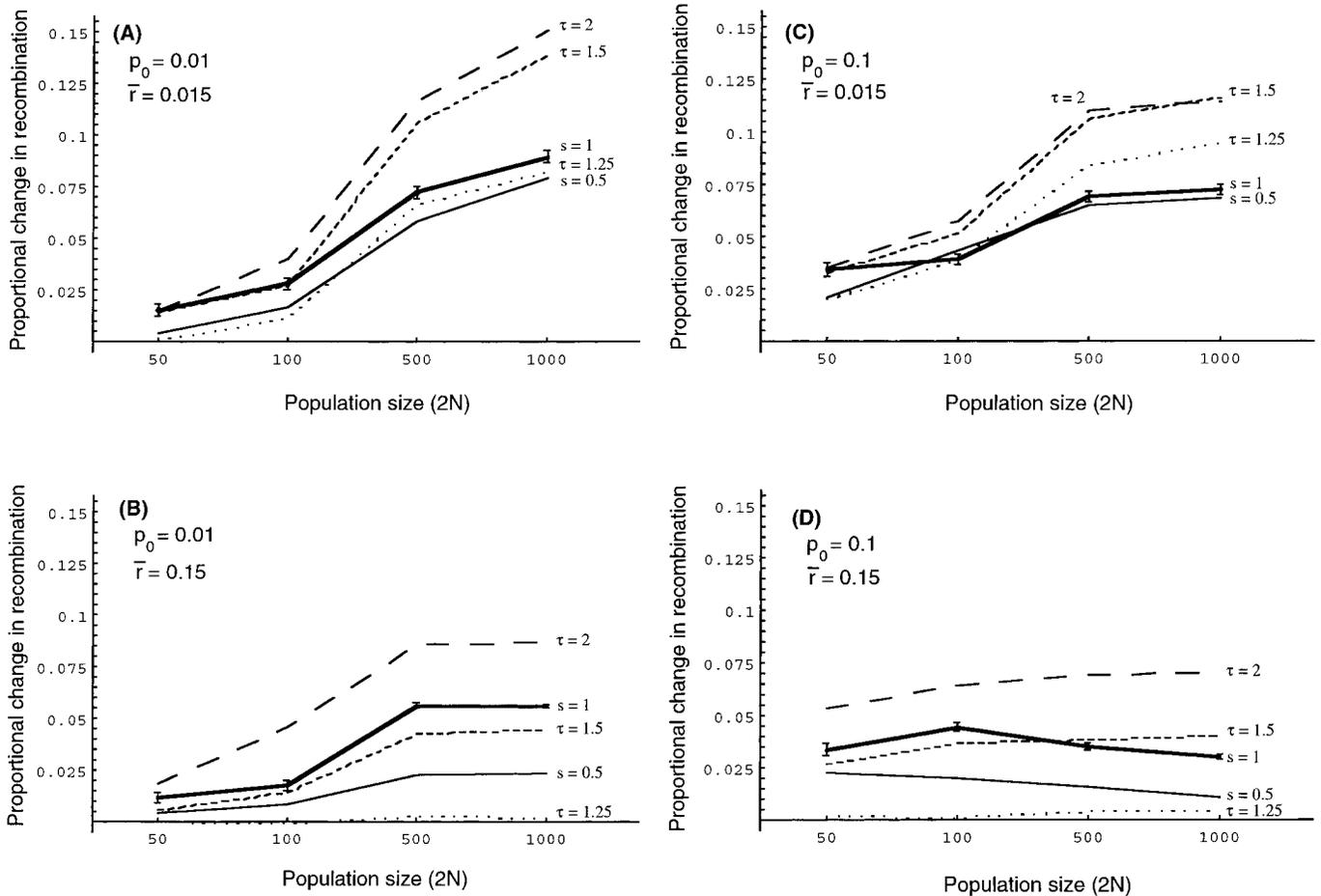
Fig. 4. The proportional change in recombination as a function of population size after 50 generations of selection in an 11-locus model. The frequency of the plus allele at the modifier locus (the last of the 11 loci) was tracked over 10^4 replicate runs with $\delta r = 1/3$. The solid curves illustrate results with strong multiplicative selection ($s = 0.5$; thin curves) and very strong multiplicative selection ($s = 1.0$; thick curves; ± 1 SE). The dashed curves illustrate results with truncation selection of varying strength: weak ($\tau = 1.25$; short dashed curves), moderate ($\tau = 1.5$; medium dashed curves), and strong ($\tau = 2$; long dashed curves), where $2N\tau$ chromosomes were examined to keep the most fit $2N$. Standard error bars are only shown for one case, but are comparable in the others. The parameters for (A–D) are as in Figure 2.

are violated, and the QLE prediction massively overestimates the change in recombination (e.g., see double dashed curves in Figs. 3A, 3C). When selection is weaker than the recombination rate, however, the QLE prediction fits the observed change reasonably well (e.g., see double dashed curves in Figs. 3B, 3D).

The results from simulations with 11 loci are presented in Figure 4, where 10 loci were under selection and the last locus on each chromosome was a modifier of recombination. In these simulations, $V_A(W)$ varied substantially with the level of truncation selection or multiplicative selection, so we must first determine which selection regimes are comparable in strength. With a few exceptions, the total change in frequency of the selected alleles over 50 generations with multiplicative selection and $s = 0.5$ fell between that observed under truncation selection with $\tau = 1.25$ and $\tau = 1.5$ (results not shown). Similarly, Δp with multiplicative selection and $s = 1$ fell between that observed with $\tau = 1.5$ and $\tau = 2$. (The exceptions were: [1] with $s = 0.5$, $p_0 = 0.1$ and $2N = 50$, Δp fell just above that observed with $\tau = 1.5$; and

[2] with $s = 1$, tight linkage, and $2N = 500$ or 1000, Δp fell between $\tau = 1.25$ and $\tau = 1.5$.) In 20 of the 32 comparisons illustrated in Figure 4, the amount of evolutionary change in recombination accomplished under multiplicative selection was intermediate between the amount accomplished by weaker truncation selection and stronger truncation selection. In 11 of the remaining 12 cases, truncation selection was more efficient at selecting for increased recombination; this was most often observed in simulations with larger populations ($2N \geq 500$ in eight of 11 of these cases). Overall, the increase in recombination under multiplicative selection divided by the increase observed under *stronger* truncation selection averaged 60%, which underestimates the true ratio of interest (where multiplicative selection is compared to truncation selection of equivalent intensity). The results suggest that the majority of the increase in recombination in these simulations can be attributed to selection in a finite population. Although truncation selection does select for increased recombination, it is neither a necessary nor the dominant evolutionary force acting on recombination in small populations.

As a check of the robustness of our results, we explored the effects of adding a second chromosome upon which selection acts. On the first chromosome, there were either two or 10 selected loci and a modifier that affected recombination rates on this chromosome. Parameters for this chromosome were the same as in Figure 2 (with two selected loci) and Figure 4 (with 10 selected loci), although only multiplicative selection with $s = 1$ and populations of size $2N = 50, 100, 500,$ and 1000 were explored. On the second chromosome, there were 10 selected loci with the same starting allele frequency (p_0), strength of selection (s), and recombination rate between adjacent loci ($\bar{r}$) as on the first chromosome, but recombination rates were not affected by the modifier. Adding a second chromosome was expected to increase the variance in reproductive success of the first chromosomes, which should cause a modest reduction in the efficacy of selection (Hill and Robertson 1966). The results were very similar to those shown in Figures 2 and 4, although the recombination rate tended to increase slightly less, on average. With two selected loci on the first chromosome, the effect on the proportional change in recombination (i.e., the effect on the y-axis of Fig. 2) was -0.00080 (ranging from -0.0071 to 0.0037 with SD of 0.0033 over 16 parameter sets). The effect with 10 selected loci on the first chromosome was -0.0027 (ranging from -0.011 to 0.0050 with SD of 0.0039). We conclude that the evolution of a locally acting modifier of recombination is only weakly affected by nonepistatic selection acting on the rest of the genome.

In general, a larger correlated increase in recombination is observed when the additive genetic variance for recombination or for fitness is higher. When the genetic variances are sufficiently high, large changes in recombination can be observed even in very small populations. To study this further, we performed two sets of simulations, first increasing the genetic variance in recombination and then increasing the genetic variance in fitness. In the first set of simulations, we added two more modifier loci to the 11-locus model, with one modifier locus at either end of the chromosome and one exactly in the middle. Thirty-two parameter sets were tested, corresponding to each combination shown in Figure 4, but with only $s = 1$ or $\tau = 2$. When the initial additive genetic variance in recombination rate was held the same as in Figure 4 by setting δr to $1/\sqrt{27}$ at each modifier locus, the proportional change in recombination that occurred over 50 generations was very similar (mean difference of 0.0031 with SD of 0.0043 over the 32 parameter sets). When the additive genetic variance in recombination was doubled by setting δr to $\sqrt{2/27}$, the evolutionary change in recombination accomplished by 50 generations of selection was also doubled (mean difference from this expectation of 0.0037 with SD of 0.0068). In the second set of simulations, we considered a population of 50 chromosomes, each 1 Morgan in length bearing 100 loci. Two of these loci (at positions 25 and 75) were modifiers of recombination, and the remainder were under multiplicative selection ($s = 0.5$). Simulations were begun with the $+$ allele at each modifier locus at a frequency of 0.5. The recombination rate would be doubled if the $+$ allele fixed at all modifier loci and would be reduced to zero if the $-$ allele fixed. After 50 generations of selection, the average length of the chromosome increased by 22% (SE = 2%; $p_0 = 0.01$) and 31% (SE = 2%; $p_0 = 0.1$) based on 1000 simulations. These increases are comparable in magnitude to many of the observed changes in recombination reported in Table 1. It should be emphasized, however, that the changes in recombination rate are very stochastic. In 27.6% ($p_0 = 0.01$) and 39.2% ($p_0 = 0.1$) of the cases, the $+$ allele fixed at both modifier loci, causing the genetic map length of the chromosome to double. Conversely, in 11.6% ($p_0 = 0.01$) and 10.1% ($p_0 = 0.1$) of the cases, the $-$ allele fixed at both modifier loci, causing a complete loss of recombination. Variability in the correlated response of recombination to selection was typical in all the simulations, even when the number of selected loci and the parameters were held constant. (Recall that the standard deviation among replicates can be obtained by multiplying the standard errors shown in the figures by the square root of the number of replicates; i.e., by 100 with 10^4 replicates and 316 with 10^5 replicates.) This must serve as a caution that any particular instance where recombination increases or decreases may be due to chance and not due to the form of selection. Thus, the great degree of variability in experimental results reported in Table 1, both in the sign and extent of recombination evolution, may be explained in part by the stochastic nature of evolution in small populations.

DISCUSSION

Changes in the genetic system are most likely to be observed during periods of strong selection. Indeed, several studies have found correlated changes in recombination rates (see Table 1) and mutation rates (LeClerc et al. 1996; Mao et al. 1997; Sniegowski et al. 1997) within strongly selected populations. But the ultimate cause of a correlated change in genetic system is difficult to identify (Hedrick 1983). Correlated changes in recombination may be caused by pleiotropy (i.e., the selected alleles directly affect recombination), inbreeding (if there is directional dominance at the modifier loci), initial disequilibria between modifier alleles and the selected alleles, initial disequilibria among the selected loci themselves, disequilibria generated by epistatic selection, or disequilibria generated by drift. In most of the selection experiments summarized in Table 1, selection was applied in only one direction, making it impossible to rule out pleiotropy or an initial association between recombination alleles and beneficial alleles (but see Korol and Iliadi 1994). In addition, many of these selection experiments were started from mixed populations derived from two or more stocks; unless these populations are allowed to undergo several generations of recombination prior to selection, such mixing heightens the risk that the major source of disequilibrium is the experimenter rather than the form of selection or drift.

Even in well-controlled experiments, with selection applied in both directions and with a minimum amount of initial disequilibria, there remain two potential explanations for an observed change in recombination. Both rely on the generation of negative disequilibria among favorable alleles, but such disequilibria could be generated by epistasis or by drift in the presence of selection. Theoretical studies of larger populations have demonstrated that both forces can select for alleles that increase recombination (Charlesworth 1993b;

Barton 1995a; Otto and Barton 1997; N. H. Barton and S. P. Otto, unpubl. ms.). That drift and selection can combine to favor recombination in small populations was also demonstrated by Felsenstein and Yokoyama (1976), who called this the Fisher-Muller theory for recombination. In a series of simulations with multiplicative directional selection, they observed that an allele that caused free recombination would fix more often than an allele causing complete linkage when both began at equal frequency. Their study, however, does not allow one to gauge the relative importance of drift and epistasis in populations subject to directional selection. In this paper, we have shown that the evolutionary advantage of recombination in small populations derives mainly from the advantage of dispelling associations between beneficial alleles and deleterious genetic backgrounds caused by random genetic drift (Figs. 2–4). Furthermore, proportional changes in recombination of the magnitude observed in artificial selection experiments (Table 1) are consistent with the Fisher-Muller theory, as long as there is ample genetic variation in fitness and in recombination. Although negative epistasis does select for recombination in large populations, the presence of epistasis generally did not make a substantial difference unless $2N$ was greater than about 500. It is important to emphasize that we deliberately chose parameters that would be most favorable to the hypothesis that epistasis favors recombination, by focusing on negative epistasis and by including the value of ϵ in the three-locus model that maximizes selection for recombination in infinite populations.

Strength of Selection

To know whether these results might apply to natural populations, we need more information about the rate of adaptive substitutions and the selection coefficients associated with them. Based on the fraction of sites exhibiting an excess of nonsynonymous substitutions, N. G. C. Smith and A. Eyre-Walker (unpubl. ms.) inferred that the rate of adaptive amino acid substitutions in *Drosophila* is approximately one every 45 years (about 450 generations). Using a similar method, Fay et al. (2001) estimated a rate of approximately one per 200 years during the divergence of humans and Old World monkeys. Alternatively, the rate of adaptive substitutions may be inferred from the reduction in genetic variability caused by hitchhiking with a beneficial allele (Maynard Smith and Haigh 1974). Schlotterer et al. (1997) observed six locus- and population-specific reductions in variability at 10 microsatellite loci surveyed in seven populations. Nurminsky (2001) argued that because a selective sweep could be detected for roughly 10^3 years, this implies one selective sweep per year, summed over a genome of 10^4 genes. This estimate is much higher than that based on amino acid divergence, perhaps because other factors are responsible for differences in variability across loci and populations or because local sweeps are much more common than specieswide substitutions.

Indirect selection on recombination modifiers also depends on the strength of selection driving these substitutions. Based on our QLE analysis (N. H. Barton and, S. P. Otto, unpubl. ms.), the per generation change in the frequency of a recombination modifier is proportional to the product of the additive variances in fitness at pairs of selected loci ($V_1 V_2$, where $V_i = s_i^2 p_i [1 - p_i]$). For a given rate of sweeps, each of which lasts on the order of $1/s_i$ generations, selection on the modifier will thus scale with the square of the selection coefficient when the QLE assumptions are met (weak selection, a weak modifier, loose linkage, and a large population size). Unfortunately, we know very little about the strength of selection in nature. Strong selection coefficients (> 0.1) have been measured within natural insect populations subject to pesticide treatment (Lenormand et al. 1999) and on alternative mimetic patterns in butterflies (Mallet et al. 1990). Endler (1986, fig. 7.2) has surveyed direct measurements of selection on visible polymorphisms and reports strong selection coefficients, which are uniformly distributed between zero and one. Of course, such strong selection may not be typical of most substitutions. However, of the 10 allozyme polymorphisms first detected in *Drosophila*, all six that have been studied in more detail show evidence of kinetic differences and most also show physiological differences (Gillespie 1991, table 1.3). This suggests that amino acid polymorphisms may often be under substantial selection within species.

Given the difficulty of directly measuring selection, a more promising approach is to draw inferences about the strength of selection from levels of molecular variation. For example, Fay and Wu (2000) identified a region of about 350 bp in the Acp26Aa gene of *D. melanogaster* with no nucleotide variability and a region twice as large in which derived variants are at high frequency. Because this gene has a rapidly evolving amino acid sequence, a recent selective sweep is the likely explanation. Because the width of the region of perturbed variability is proportional to the selection coefficient, Fay and Wu (2000) estimate that the selection acting must be at least $s \approx 0.004$. Using the observed reduction in nucleotide variability in regions of low recombination in *D. melanogaster*, Stephan (1995) estimated the intensity of selective sweeps, $2Ns\Lambda$ (twice the product of effective population size, selection coefficient, and rate of sweeps), to be 6.6×10^{-8} per nucleotide per year, assuming that the reduced variability is entirely due to selective sweeps. Some of the reduction must in fact be due to linked deleterious mutations (background selection); however, Kim and Stephan (2000) argue that selective sweeps are the dominant cause of reduced variability in regions of low recombination. By combining Stephan's (1995) figure with their estimate of the rate of sweeps, A. Eyre-Walker (pers. comm.) estimates that $2Ns$ lies in the range of 500–5000. Dividing by the average effective population size of *D. melanogaster* and *D. simulans*, estimated from nucleotide variability, gives estimates of s between 10^{-3} and 2×10^{-4}. These estimates are very much smaller than those suggested by direct measurements. One explanation might be that in spatially structured populations, substitutions occur more slowly than in a panmictic population, thereby reducing their effect on linked loci (Barton 2000). Another potential concern is that recombination may itself be mutagenic and may contribute to differences in genetic variability between regions of low and high recombination (Marais et al. 2001).

These various estimates of the rate of adaptive substitutions

are quite uncertain; estimates of the strength of selection are even more approximate. Moreover, the effect of positive selection on the evolution of recombination depends strongly on variability in rates and strengths of selection. Even if beneficial alleles arise rarely and tend, on average, to be weakly selected, occasional bouts of strong selection associated with a change in the environment or colonization of a new environment may have a disproportionate influence on the evolution of recombination.

Future Directions

To further understand the forces acting on recombination, selection experiments are needed with replicate populations whose sizes vary over orders of magnitude. In such an experiment, every effort should be made to reduce initial disequilibria among loci to ensure that the evolution of recombination is not trivially due to initial disequilibria among selected loci or among modifiers and selected loci (Hedrick 1983). Ideally, replicate populations should originate from different stocks to reduce the chance that they all start with the same genetic associations (Zhuchenko et al. 1988), they should be genetically variable, and they should have been maintained under control conditions for several generations. Preferably, directional selection should be produced by exposure to a novel environment rather than by truncation selection to ensure that the form of epistasis arises naturally and is not simply a reflection of the experimental methods. To control for the possibility that the correlated response to selection is due to pleiotropy, selection should be performed in both upward and downward directions, or crosses should be performed at the end of the experiment to determine the extent to which modifiers of recombination segregate with the genetic factors responsible for the direct response to selection. Finally, the appropriate model organism should be amenable to study over a wide range of effective population sizes, such as 10^2, 10^3, 10^4, and 10^5. It may be easier to work with organisms that are not obligately sexual, in that the frequency of sex could be scored rather than the recombination rate among markers or the number of chiasmata. If the advantage of sex and recombination comes primarily from bringing together beneficial alleles that happen, by chance, to occur in different individuals in small populations, then correlated increases in the frequency of sex or recombination should diminish in experiments with very large populations (e.g., 10^5). However, if the correlated increase in recombination persists in large populations (e.g., 10^4) and is insensitive to further increases in population size (e.g., to 10^5), the presence of negative epistasis may be inferred. Beyond distinguishing between the two mechanisms, such an experiment is necessary to know whether rapid and substantial changes in the frequency of sex and recombination may be seen in large populations, given that the majority of the experiments (and the ones with the clearest positive results) have been performed on small populations.

On the theoretical side, work needs to be done to determine whether the Fisher-Muller mechanism ever leads to substantial selection for increased levels of sex and recombination in populations of large size. One possibility is that increased recombination rates could evolve by the Fisher-Muller mech-

anism if selection acts primarily on local demes of small size, even if the total population size is large. If deme sizes are typically on the order of 1000 or less, the Fisher-Muller mechanism could be effective, as was seen in this study. When the form of selection varies among demes, however, migration also generates disequilibria by importing chromosomes bearing alleles that are more common elsewhere (Lenormand and Otto 2000). It remains to be determined whether the disequilibria generated by local genetic drift could be large enough both to overwhelm the disequilibria generated by migration and epistasis and to account for the evolution of recombination rates in large populations.

LITERATURE CITED

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

———. 1995b. Linkage and the limits to natural selection. Genetics 140:821–841.

———. 2000. Genetic hitchhiking. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 355:1553–1562.

Barton, N. H., and B. Charlesworth. 1998. Why sex and recombination? Science 281:1986–1990.

Burt, A. 2000. Perspective. Sex, recombination, and the efficacy of selection: Was Weismann right? Evolution 54:337–351.

Burt, A., and G. Bell. 1987. Mammalian chiasma frequencies as a test of two theories of recombination. Nature 326:803–805.

Charlesworth, B. 1993a. Directional selection and the evolution of sex and recombination. Genet. Res. 61:205–224.

———. 1993b. The evolution of sex and recombination in a varying environment. J. Hered. 84:345–350.

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

Endler, J. A. 1986. Natural selection in the wild. Princeton Univ. Press, Princeton, NJ.

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

Fay, J. C., and C.-I. Wu. 2000. Hitchhiking under positive Darwinian selection. Genetics 155:1405–1413.

Fay, J. C., G. J. Wyckoff, and C.-I. Wu. 2001. Positive and negative selection on the human genome. Genetics 158:1227–1234.

Feldman, M. W. 1972. Selection for linkage modification. I. Random mating populations. Theor. Popul. Biol. 21:430–439.

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

Felsenstein, J., and S. Yokoyama. 1976. The evolutionary advantage of recombination. II. Individual selection for recombination. Genetics 83:845–859.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford Univ. Press, Oxford, U.K.

Flexon, P. B., and C. F. Rodell. 1982. Genetic recombination and directional selection for DDT resistance in Drosophila melanogaster. Nature 298:672–674.

Gillespie, J. H. 1991. The causes of molecular evolution. Oxford Univ. Press, Oxford, U.K.

Gorlov, I., L. Schuler, L. Bunger, and P. Borodin. 1992. Chiasma frequency in strains of mice selected for litter size and for high body weight. Theor. Appl. Genet. 84:640–642.

Gorodetskii, V. P., A. A. Zhuchenko, and A. B. Korol. 1990. Efficiency of feedback selection for recombination in *Drosophila*. Genetika 26:1942–1952.

Haldane, J. B. S. 1990. The causes of evolution. Princeton Univ. Press, Princeton, NJ.

Harinarayana, G., and B. R. Murty. 1971. Cytological regulation of recombination in *Pennisetum* and *Brassica*. Cytologia 36: 435–448.

Hedrick, P. W. 1983. Recombination and directional selection. Nature 302:727.

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

Kim, Y., and W. Stephan. 2000. Joint effects of genetic hitchhiking and background selection on neutral variation. Genetics 155: 1415–1427.

Kondrashov, A. S., and L. Y. Yampolsky. 1996. Evolution of amphimixis and recombination under fluctuating selection in one and many traits. Genet. Res. Camb. 68:165–173.

Korol, A. B., and K. G. Iliadi. 1994. Increased recombination frequencies resulting from directional selection for geotaxis in *Drosophila*. Heredity 72:64–68.

Korol, A. B., I. A. Preygel, and S. I. Preygel. 1994. Recombination variability and evolution: algorithms of estimation and population-genetic models. Chapman and Hall, London.

LeClerc, J. E., B. Li, W. L. Payne, and T. A. Cebula. 1996. High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. Science 274:1208–1211.

Lenormand, T., and S. P. Otto. 2000. The evolution of recombination in a heterogeneous environment. Genetics 156:423–438.

Lenormand, T., D. Bourguet, T. Guillemaud, and M. Raymond. 1999. Tracking the evolution of insecticide resistance in the mosquito *Culex pipiens*. Nature 400:861–864.

Lobashev, M. E., V. V. Ponomarenko, G. G. Polyanskaya, and R. I. Tsapygina. 1973. The role of the nervous system in the regulation of various genetic and cytogenetic processes. J. Evol. Biochem. Physiol. 9:398–405.

Mallet, J., N. Barton, G. Lamas, J. Santisteban, M. Muedas, and H. Eeley. 1990. Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in Heliconius hybrid zones. Genetics 124:921–936.

Mao, E. F., L. Lane, J. Lee, and J. H. Miller. 1997. Proliferation of mutators in a cell population. J. Bacteriol. 179:417–422.

Marais, G., D. Mouchiroud, and L. Duret. 2001. Does recombination improve selection on codon usage? Lessons from nematode and fly complete genomes. Proc. Natl. Acad. Sci. USA 98: 5688–5692.

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

———. 1980. Selection for recombination in a polygenic model. Genet. Res. 35:269–277.

———. 1988. Selection for recombination in a polygenic model: the mechanism. Genet. Res. 51:59–63.

Maynard Smith, J., and J. Haigh. 1974. The hitch-hiking effect of a favorable gene. Genet. Res. Camb. 23:23–35.

Michod, R. E., and B. R. Levin. 1988. The evolution of sex. Sinauer Associates, Sunderland, MA.

Morgan, T. H. 1913. Heredity and sex. Columbia Univ. Press, New York.

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

Nurminsky, D. I. 2001. Genes in sweeping competition. Cell. Mol. Life Sci. 58:125–134.

Otto, S. P., and N. H. Barton. 1997. The evolution of recombination: removing the limits to natural selection. Genetics 147:879–906.

Otto, S. P., and Y. Michalakis. 1998. The evolution of recombination in changing environments. Trends Ecol. Evol. 13: 145–151.

Palenzona, D. L., M. Mochi, and R. Alicchio. 1972. Recombination rate in selected lines of *Drosophila melanogaster* Meig. Monitore Zool. Ital. 6:231–237.

Schlotterer, C., C. Vogl, and D. Tautz. 1997. Polymorphism and locus-specific effects on polymorphism at microsatellite loci in natural *Drosophila melanogaster* populations. Genetics 146: 309–320.

Sniegowski, P. D., P. J. Gerrish, and R. E. Lenski. 1997. Evolution of high mutation rates in experimental populations of *E. coli*. Nature 387:703–705.

Stephan, W. 1995. An improved method for estimating the rate of fixation of favorable mutations based on DNA polymorphism data. Mol. Biol. Evol. 12:959–962.

West, S. A., C. M. Lively, and A. F. Read. 1999. A pluralist approach to sex and recombination. J. Evol. Biol. 12:1003–1012.

Wolf, H. G., K. Wöhrmann, and J. Tomuik. 1987. Experimental evidence for the adaptive value of sexual reproduction. Genetica 72:151–159.

Zhuchenko, A. A., A. B. Korol, and L. P. Kovtyukh. 1985. Change of crossing-over frequency in *Drosophila* during selection for resistance to temperature fluctuations. Genetica 67:73–78.

Zhuchenko, A. A., A. B. Korol, and V. P. Gorodetskii. 1988. Experimental modeling of the microevolution of recombination. Dokl. Akad. Nauk SSSR 302:452–456.

Corresponding Editor: A. Caballero