The Evolution of Recombination: Removing the Limits to Natural Selection

Sarah Perin Otto* and Nick H. Barton†

**Department of Zoology, University of British Columbia, Vancouver, BC, Canada V6T 1Z4 and †Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom*

Manuscript received September 2, 1996
Accepted for publication June 12, 1997

ABSTRACT

One of the oldest hypotheses for the advantage of recombination is that recombination allows beneficial mutations that arise in different individuals to be placed together on the same chromosome. Unless recombination occurs, one of the beneficial alleles is doomed to extinction, slowing the rate at which adaptive mutations are incorporated within a population. We model the effects of a modifier of recombination on the fixation probability of beneficial mutations when beneficial alleles are segregating at other loci. We find that modifier alleles that increase recombination do increase the fixation probability of beneficial mutants and subsequently hitchhike along as the mutants rise in frequency. The strength of selection favoring a modifier that increases recombination is proportional to $\lambda^2 S \delta r / r$ when linkage is tight and $\lambda^2 S^3 \delta r / N$ when linkage is loose, where λ is the beneficial mutation rate per genome per generation throughout a population of size N, S is the average mutant effect, r is the average recombination rate, and δr is the amount that recombination is modified. We conclude that selection for recombination will be substantial only if there is tight linkage within the genome or if many loci are subject to directional selection as during periods of rapid evolutionary change.

ALTHOUGH sexual reproduction and recombination are widespread among eukaryotes, the forces that are primarily responsible for the evolution of recombination continue to be hotly debated [for many of the arguments see articles in MICHOD and LEVIN 1988 and in the special issue of the *Journal of Heredity* devoted to this topic (1993, Vol. 84, no. 5)]. One of the oldest hypotheses for the advantage of sex and recombination is that it allows beneficial mutations, initially carried by different individuals, to be combined together into the same individual (FISHER 1930; MULLER 1932; see also MORGAN 1913, p. 14). In other words, mutations that arise on a poor genetic background can recombine onto a better genetic background. This uncouples the fate of a mutation from the fate of the rest of the genome in which the mutation appears. As a consequence, recombination tends to increase the probability of fixation of beneficial mutations (BARTON 1995b). In this article, we demonstrate that this stochastic advantage to recombination favors the evolution of increased recombination rates at modifier loci even in the absence of selective interactions (epistasis) among loci. Before we describe our model further, we begin by reviewing other hypotheses for the advantage of recombination.

The various hypotheses for the advantage of genetic recombination fall naturally into two groups (*cf.* MAYNARD SMITH 1978, p. 73; KONDRASHOV 1993): (1) recombination confers an immediate benefit (physiological hypotheses) and (2) recombination generates variation that may be beneficial in the face of selection (generative hypotheses). Experimental evidence indicates that neither a physiological nor a generative explanation is wholly satisfactory and that the evolution of recombination has been governed by elements of both types of explanation. Understanding this evidence is critical if we are to assess the merits of any particular hypothesis.

Physiological hypotheses: Recombination plays a critical physiological and mechanistic role in the life cycle of the cell. One such role is in the repair of DNA damage, which has led several authors to argue that DNA repair is a major function of recombination and the primary one in bacteria (BERNSTEIN *et al.* 1988; MICHOD and LEVIN 1988; COX 1993). The repair hypothesis for the evolution of recombination is supported by the finding that many proteins that were originally identified as required for recombination are also critical for the repair of DNA lesions, cross-links and double strand breaks [*e.g.*, the RecA, RecE, RecJ, and RecQ proteins in *Escherichia coli* (COX 1993; KUSANO *et al.* 1994) and RAD51, RAD52, and RAD54 proteins in *Saccharomyces cerevisiae* (BASILE *et al.* 1992)]. Mutants at these loci recombine at a lower frequency and often fail to replicate in the presence of DNA damage. Under this repair hypothesis, the advantage of recombination derives from the immediate benefit of ensuring cell survival in the face of DNA damage.

Another important physiological role of recombination is the stabilization of homologous chromosomes during meiosis in sexual eukaryotes (BAKER *et al.* 1976;

Corresponding author: Sarah Perin Otto, Department of Zoology, 6270 University Blvd., University of British Columbia, Vancouver, BC, Canada V6T 1Z4. E-mail: otto@zoology.ubc.ca

HAWLEY *et al.* 1988). Chiasmata (the cytological manifestation of genetic exchange between chromosomes) serve to hold non-sister chromosome pairs together during meiosis until they migrate apart at anaphase I. Typically, there is at least one chiasma per chromosome pair per meiosis, with the frequency of achiasmate chromosomes being much less than expected based on a Poisson distribution (KNAPP 1960; HAWLEY 1988; BURT *et al.* 1991). In studies from a variety of organisms including *Drosophila melanogaster* (BAKER *et al.* 1976), *S. cerevisiae* (ENGEBRECHT *et al.* 1990), *Caenorhabditis elegans* (VILLENUEVE 1994), and humans (KOEHLER *et al.* 1996), the frequency of achiasmate chromosomes has been found to correlate positively with abnormal segregation and aneuploidy. Interestingly, an excessive number of crossover events per chromosome may also impair segregation (MERRIAM and FROST 1964; KOEHLER *et al.* 1996), presumably because of the mechanical difficulties in separating highly intertwined homologues during meiosis. These findings are most consistent with the view that the need for proper disjunction during meiosis places selective constraints on the frequency of chiasmata. These selective constraints are often dismissed by evolutionary biologists for two reasons (*e.g.*, MAYNARD SMITH 1978, p. 74; CHARLESWORTH *et al.* 1985). The first is that a number of species, including *D. melanogaster,* have dispensed with chiasmata altogether in one sex (BURT *et al.* 1991). A wide variety of alternative segregation systems are known to help ensure proper disjunction of chromosomes during meiosis, including the distributive system in Drosophila females and pairing-site-mediated segregation in Drosophila males (HAWLEY 1988; HAWLEY and THEURKAUF 1993). These systems make it possible for a species to lose chiasmata in one sex without dire fitness costs. Nevertheless, these "back-up" systems are often limited in distribution to a particular group of species and are often not failsafe. For example, aneuploidy is more frequent in the absence of cross-overs in Drosophila females despite the existence of the distributive system (which actually involves two separate achiasmate segregation mechanisms; HAWLEY and THEURKAUF 1993). Therefore, even though a species may evolve other means to ensure proper disjunction, "genetic recombination, or crossing-over, is the primary means by which the disjunction of homologous chromosomes at meiosis I is ensured" (KOEHLER *et al.* 1996, p. 1495). The second reason why evolutionary biologists have dismissed the notion of selective constraints on the evolution of recombination is that, even if chiasmata were necessary for segregation, it is presumed that their position is free to vary. For example, if no recombination were advantageous, selection could favor crossover mechanisms that preferentially target the ends of chromosomes, which would ensure segregation but would entail little recombination. There is evidence, however, that the frequency of aneuploidy also increases when chiasmata occur at the ends of chromosomes, perhaps because the tension provided by terminal chiasmata may not be sufficient to prevent premature separation of chromosomes during meisois (KOEHLER *et al.* 1996). We therefore conclude that, in most species, the need for proper segregation acts as a selective force constraining the position as well as the number of chiasmata. It is possible for other selective forces to be sufficiently strong that evolution may overcome these constraints, but it remains to be determined how often this occurs.

Generative hypotheses: The above discussion suggests that the physiological functions of recombination might completely govern its evolution, with the generation of genetic variability being a mere byproduct (COX 1993). The strongest evidence that generating genetic variability must also play a role in the evolution of recombination rates comes from artificial selection experiments. In a number of experiments in a variety of organisms, it has been observed that recombination rates increase after a period of direct selection upon traits not associated with recombination (reviewed in KOROL and ILIADI 1994). For example, KOROL and ILIADI (1994) selected both positively and negatively for geotaxis in *D. melanogaster* and observed substantial increases in recombination. Across all regions studied, recombination rates rose from 217.0 cM in the control line to 295.9 cM in the geo^+ and 283.4 cM in the geo^- line over a period of 50 generations. KOROL and ILIADI concluded that "directional selection for any trait for many generations can play an important role in *rec*-system evolution." Such results are difficult to reconcile with any of the physiological explanations for recombination. Instead, they suggest that recombination does evolve, at least in part, to generate genetic variability in the face of selection.

A number of theories for the advantage of recombination have been proposed that focus on the generative role of recombination (see reviews by KONDRASHOV 1993 and FELDMAN *et al.* 1996). By changing the array of offspring genotypes expected upon reproduction, recombination affects the fitness distribution among descendants of an individual, leading to indirect selection on recombination that can cause recombination rates to evolve. This process has been studied using modifier models that track the dynamics of loci that alter recombination rates (NEI 1967; FELDMAN 1972). Results from these models have demonstrated that producing recombinant offspring may favor increased recombination, but need not. For instance, in a constant environment without mutation, recombination is selected against (ALTENBERG and FELDMAN 1987). In a fluctuating environment, increased recombination rates can evolve, but the conditions that favor increased recombination are fairly restrictive (CHARLESWORTH 1976, 1990; BARTON 1995a). Finally, with selection either against deleterious mutations or for beneficial mu-

tations, recombination can be favored if there is negative epistasis among mutations, that is, if the fitness of an individual with multiple deleterious (or beneficial) mutations is lower than expected from the product of their individual effects (FELDMAN *et al.* 1980; KONDRASHOV 1982; CHARLESWORTH 1990; BARTON 1995a; OTTO and FELDMAN 1997). This last hypothesis is currently favored by many evolutionary biologists. The main problem with it is that there is limited evidence for the existence of negative epistasis [coming largely from theoretical considerations and from an experimental study by MUKAI (1969) in *D. melanogaster*]. A further problem is that the advantage of recombination is lessened and can even disappear if loci vary in their interactions, with some loci interacting strongly and others weakly or not at all (OTTO and FELDMAN 1997).

All of the above modifier models have, however, ignored stochastic effects such as sampling fluctuations experienced by new mutations. Even in large populations, the effects of random sampling are substantial when mutations first appear and recombination can play an important role in this process. It has been shown that a new beneficial mutation is more likely to fix if it is unlinked to other selected loci; the higher the recombination rate, the higher the probability of fixation (HILL and ROBERTSON 1966; FELSENSTEIN 1974; BARTON 1995b). It is this effect that we explore in this paper. We demonstrate that increased recombination rates can evolve at a modifier locus, even in the absence of epistatic interactions among genes, as a result of the increased fixation probability of beneficial mutations. This result may help explain the increased recombination rates observed in the directional selection experiments cited by KOROL and ILIADI. Although these authors favored the negative epistasis hypothesis to explain the advantage of recombination after a period of strong directional selection, it seems unlikely to us that, in each example cited by KOROL and ILIADI, there is negative epistasis among the selected set of loci. The hypothesis studied in this paper, that recombination increases the rate of fixation of beneficial mutations and thereby the rate of adaptation, is an alternative explanation for these experimental results. It is more parsimonious in that the random sampling of individual mutations necessarily leads to interference among linked loci in their response to selection, whereas the deterministic alternatives require that there be negative epistasis (BARTON 1995a).

This hypothesis has a long history in evolutionary biology, a history that is worth summarizing since our results touch on some recurring themes in the literature (for further details see the review by FELSENSTEIN 1974).

Historical perspective: Early arguments for the evolution of recombination (MORGAN 1913; FISHER 1930; MULLER 1932) focused on the ability of recombination to place beneficial mutations together on the same chromosome.

> For, unless advantageous mutations occur so seldom that each has had time to become predominant before the next appears, they can only come to be simultaneously in the same gamete by means of recombination.
>
> —R. A. FISHER (1930, p. 104)

This hypothesis was later modeled by CROW and KIMURA (1965). They argued that, in asexual lineages, mutations would only survive if they occur within the single lineage that is destined, in the long run, to leave descendants. In sexual lineages, however, all mutations could potentially survive. According to their calculations, the relative rate of incorporation of new mutations could be several orders of magnitude larger with recombination than without recombination. This advantage was verified by HILL and ROBERTSON (1966), who showed by simulation that new mutations were more likely to fix in organisms with recombination. A strong counter-argument was immediately produced by MAYNARD SMITH (1968). He noted that, in a deterministic model with multiplicative selection and with no initial disequilibrium, the frequency of the various haplotypes is independent of the recombination fraction in all future generations. Consequently, recombination neither hastens nor impedes the rate of adaptation. In his model, recombination simply did not affect the probability that beneficial alleles were found together in the same genome.

Does recombination provide the immense advantage argued by CROW and KIMURA (1965) or no necessary advantage as argued by MAYNARD SMITH (1968)? The discrepancy between these arguments is subtle. MAYNARD SMITH maintained that it was unreasonable to assume, as CROW and KIMURA did, that mutations at a locus are unique events, occurring in a single individual within a population. Instead, he assumed that mutations are produced in a population at a certain rate, a rate that is independent of the selective advantage of the mutation and its linkage relationships with other loci. MAYNARD SMITH's assumption of a constant influx of mutations is, however, also untenable for finite populations. Complex mutations (duplications, rearrangements) may only occur once in the history of a population. Specific point mutations, while not so rare, are still uncommon. The mutation rate in mammals is $\approx 4.6 \times 10^{-9}$ per site per year based on the rate of synonymous substitutions (LI and GRAUR 1991, p. 71). In addition, most of these mutations are lost soon after they arise. Therefore, factors such as recombination, which increase the fixation probability of mutations, can reduce the waiting time for a successful mutation, even when such mutations are recurrent.

HALDANE (1927) was the first to study the fixation probability, P, of a particular mutant allele in a large but not infinite population. With s measuring the selective advantage of a mutant allele at a locus and ignoring all other loci within the genome, he showed that P is

the positive root of the equation $1 - P = e^{-(1+s)P}$, which for weak selection gives $P \approx 2s$. With multiple loci and multiplicative or additive selection, the probability of fixation of a beneficial mutation is always reduced when other substitutions are in progress, especially when the loci are linked (HILL and ROBERTSON 1966; FELSENSTEIN 1974; BARTON 1995b). Dynamically, this occurs because randomly generated linkage disequilibrium has the effect, on average, of reducing the response of allele frequencies to selection (KARLIN 1973). For instance, it is straightforward to show, using a simple two-locus haploid model, that variance in disequilibrium always decreases the expected change in allele frequency of a selected locus, even when the mean disequilibrium is zero. A single beneficial mutation appearing at one locus on a random genetic background with respect to a second locus has an expected disequilibrium of zero but a positive variance. Averaged over all genetic backgrounds, the randomly generated linkage disequilibria decrease the rate at which the new beneficial mutation rises in frequency and therefore reduce the chance that the allele will survive loss due to random sampling. Recombination, by reducing genetic associations, reduces the interference between loci in their response to selection, bringing the probability of fixation back toward $2s$ (BARTON 1995b). An alternative view of this effect focuses on changes in the effective population size; selection acting on a number of loci causes random fluctuations analogous to those due to random drift and thereby reduces the power of selection over drift at any particular locus (HILL and ROBERTSON 1966; BARTON 1995b). The magnitude of the advantage to recombination from reducing interference among loci was, however, largely overestimated by CROW and KIMURA, who assumed that recombination would enable the fixation of all beneficial mutations (see BODMER 1970). BARTON (1995b) recently derived the fixation probability of a beneficial mutation at one locus, an arbitrary distance from a second selected locus. Using a similar approach, we follow the dynamics of a modifier locus whose alleles alter the rate of recombination between selected loci and consequently affect the fixation probability of beneficial mutations. This work is the first analytical study to assess whether recombination may have evolved to increase the fixation probability of adaptive mutations and hence remove the limits that are placed on natural selection by the initial genetic associations of beneficial mutations.

MODEL

We consider a three locus model in which two loci (J and K) are under directional selection and a third locus (M) modifies recombination rates between the three loci (notation is summarized in Table 1). We will assume that the gene order is MJK, although the analysis applies more generally. Directional selection favors allele k_1 over k_0 at locus K (with selection coefficient S) and allele j_1 over j_0 (with selection coefficient s). The frequency of the favored allele 1 at a locus (say K) is measured by p_k, while the alternative allele 0 is measured by $q_k = 1 - p_k$. Selection acts either at the haploid or diploid stage and is assumed to be multiplicative both within and across loci. The alleles m_1 and m_0 at the modifier locus control the rates of recombination among the loci. Denote the rate of recombination between loci M and J as r_1 and between J and K as r_2, where these are average rates observed over all modifier genotypes initially present within the population. In a given individual, we assume that replacing an m_0 allele with an m_1 allele increases the recombination rate by a small amount (δr_1 between M and J and δr_2 between J and K) as shown in Table 2 (see also BARTON 1995a).

We make two key assumptions. First, we assume that the modifier is weak ($\delta r \ll r$). Then, though the average rate of recombination will change over time as the modifier alleles change in frequency, this effect is negligible: the allele frequency change at locus M will on average be proportional to δr and therefore the change in average recombination rate will be order $(\delta r)^2$. Second, we assume that the population is very large ($2Ns \gg 1$), so that the fate of a new favorable mutation is decided while it is still rare. (The evolution of recombination in small populations will be examined in a subsequent article.)

TABLE 1

Summary of notation

p_m, p_j, p_k	Frequency of allele 1 at loci M, J, and K, respectively
q_m, q_j, q_k	Frequency of allele 0 at loci M, J, and K, respectively
s	Selective advantage of the rare beneficial mutation at locus J
S	Selective advantage of the beneficial allele segregating at locus K
θ	Scaled advantage of the rare allele ($\theta = s/S$)
r_1 (r_2)	Recombination between loci MJ (JK)
δr_1 (δr_2)	Change in r_1 (r_2) due to the modifier
ρ_1 (ρ_2)	Scaled recombination rates ($\rho_1 = r_1/S$; $\rho_2 = r_2/S$)
P_{mk}	Fixation probability of a single rare allele when linked with the i^{th} allele at the modifier locus and the k^{th} allele at the second selected locus
P_k	Average fixation probability of a single rare allele on genetic background k
δP_k	Change in P_k due to a modifier
Π	Average fixation probability across all backgrounds scaled by the expectation of $2s$. $\Pi = (p_k P_1 + q_k P_0)/(2s)$.
Δ	Difference in fixation probability between genetic backgrounds scaled by $2s$. $\Delta = (P_1 - P_0)/(2s)$.
$\delta\Pi$	Effect of modifier on average fixation probability scaled by $2s\,\delta r_2/S$. $\delta\,\Pi = (p_k\,\delta P_1 + q_k\,\delta P_0)/(2s\,\delta r_2/S)$.
$\delta\Delta$	Effect of modifier on difference in fixation probability between genetic backgrounds scaled by $2s\,\delta r_2/S$. $\delta\Delta = (\delta P_1 - \delta P_0)/(2s\,\delta r_2/S)$.

TABLE 2

Recombination rates with a modifier

Genotype	Frequency	Recombination rate	
		Between M and J	Between J and K
m_1m_1	p_m^2	$r_1 - \delta r_1(p_m - q_m) + \delta r_1 = r_1 + 2\delta r_1 q_m$	$r_2 + 2\,\delta r_2\, q_m$
m_0m_1	$2p_mq_m$	$r_1 - \delta r_1(p_m - q_m)$	$r_2 - \delta r_2(p_m - q_m)$
m_0m_0	q_m^2	$r_1 - \delta r_1(p_m - q_m) - \delta r_1 = r_1 - 2\delta r_1 p_m$	$r_2 - 2\,\delta r_2\, p_m$
Average		r_1	r_2

When there is a polymorphism at a locus, K, with allele k_1 sweeping through the population, it has been shown that the probability of fixation of new beneficial mutations at a second locus, J, is higher when recombination between the selected loci is frequent (BARTON 1995b). Consequently, the mutant allele j_1 will be more likely to fix when it arises with the high-recombination allele, m_1, than when it arises with m_0. Thereafter, if j_1 rises to fixation, it will tend to carry along with it the m_1 allele. While the change in allele frequency at a neutral locus in response to selection at another locus is known as "hitchhiking," this case differs from most examples of hitchhiking in that it is not neutral. Instead we expect a change in the allele frequencies at the M locus since m_1 increases the fixation probability of j_1 and therefore is more likely to experience a hitch than the m_0 allele. This is the effect that we wish to quantify. To start, we must estimate the probability of fixation of allele j_1 when it first appears within the population on each of the four backgrounds made possible by polymorphism at both the modifier locus (M) and the selected locus (K).

We will assume that when the new beneficial mutation arises, there is no linkage disequilibrium between the modifier of recombination and the locus (K) that is already polymorphic. Furthermore, we will assume that the mutation arises at random on one of the four genetic backgrounds according to their frequencies in the population at that time. Therefore, on average, the modifier locus is initially at linkage equilibrium with the other loci. Denote the probability of fixation of the mutant allele j_1 on a particular genetic background by P_{mk}, where m denotes the allele at the modifier locus (1 for m_1 and 0 for m_0) and k denotes the allele at the other selected locus (1 for k_1 and 0 for k_0).

First consider the simple case of one selected locus. The fixation probability, $P[t]$, of a single copy of the j_1 allele is measured before selection in generation t. Reproduction then occurs such that the number of offspring alleles expected from a j_1 allele is Poisson distributed with mean $1 + s$ (fertility selection). The probability of eventual fixation of a single copy of the beneficial mutation in the next generation, $P[t+1]$, is then related to $P[t]$ by

$$1 - P[t] = \sum_{j=0}^{\infty} e^{-(1+s)} \frac{(1+s)^j}{j!} (1 - P[t+1])^j,$$

since the probability that the parental allele is eventually lost is equal to the probability that all the offspring alleles are eventually lost. Evaluating the sum, we have

$$1 - P[t] = e^{-(1+s)P[t+1]} \approx 1 - (1+s)P[t+1] + \frac{P[t+1]^2}{2} + o(s^2),$$

since P is order s, which is small (HALDANE 1927; BARTON 1995b). In continuous time, the fixation probability of a single beneficial mutation changes over time as

$$\frac{dP}{dt} \approx -sP + \frac{P^2}{2}.$$

In the simplest one-locus model, changes over time in the genetic background are ignored and selection is assumed to remain constant, in which case $P[t+1] = P[t]$ and $dP/dt = 0$, and we obtain the classic result that the probability of fixation of a new beneficial allele is approximately $2s$ (HALDANE 1927).

In an analogous manner, the probability of fixation of a mutation can be calculated in a multi-locus context by taking into account changes that occur in the genetic background in which a mutation is found (more details may be found in BARTON 1995b). In the three locus model described above, the exact iteration for the probability of fixation of a single copy of j_1 initially linked to m_1 and k_1 is

$$1 - P_{11}[t] = \exp\left\{-(1+s)\frac{1+S}{1+p_kS} P_{11}^*[t+1]\right\}. \quad (1)$$

$(1+s)(1+S)/(1+p_kS)$ is the fitness of the $m_1j_1k_1$ haplotype relative to the mean fitness in the population at the time ($=1 + p_kS$). $P_{11}^*[t+1]$ is the expected fixation probability of a single offspring copy of j_1 taking into account all possible ways in which its genetic background could change. By assuming random pairing with the other haplotypes in the population,

$$\begin{aligned} P_{11}^*[t+1] = {} & (p_mp_k^A + p_mq_k^A(1 - (r_2 + 2\delta r_2 q_m)) \\ & + q_mp_k^A(1 - r_1) + q_mq_k^A(1 - r_1) \\ & \times (1 - (r_2 - \delta r_2(p_m - q_m)))) P_{11}[t+1] \\ & + (p_mq_k^A(r_2 + 2\,\delta r_2\, q_m) + q_mq_k^A(1 - r_1) \\ & \times (r_2 - \delta r_2(p_m - q_m))) P_{10}[t+1] + (q_mp_k^A r_1 \\ & + q_mq_k^A r_1(1 - (r_2 - \delta r_2(p_m - q_m)))) P_{01}[t+1] \\ & + (q_mq_k^A r_1(r_2 - \delta r_2(p_m - q_m))) P_{00}[t+1]. \quad (2) \end{aligned}$$

This formula applies whether the population is diploid, in which case the pairing has occurred before fertility selection, or haploid in which case the pairing will occur after fertility selection. In either case, the probability of pairing with an m_1 allele is p_m and with an m_0 allele is $q_m = 1 - p_m$, but the probability of pairing with a k_1 allele must take into account the fertility advantage of k_1 in partner haplotypes, since haplotypes bearing the k_1 allele are expected to produce more offspring (by a factor $1 + S$) than haplotypes bearing the k_0 allele. That is, we must weight the probability of pairing by the expected number of offspring produced by each pair. This is equivalent to setting the probability of pairing with k_1 to the frequency of k_1 after selection, $p_k^A = (1 + S)p_k/(1 + p_kS)$, and the probability of pairing with k_0 to $q_k^A = 1 - p_k^A$.

The first part of (2) multiplying $P_{11}[t + 1]$ gives the probability that no change in background occurs. The second part multiplying $P_{10}[t + 1]$ gives the probability that j_1 becomes linked with m_1 and k_0 in the next generation, etc. For ease of writing, we have redefined r_1 to equal the rate of recombination between M and J in m_0m_1 heterozygotes (that is, $r_1 - \delta r_1(p_m - q_m) \to r_1$). This redefinition is reasonable since only the recombination rate between M and J in heterozygotes enters into the equation. The recombination rate between M and J in m_0m_0 and m_1m_1 individuals is immaterial since such a recombination event will not lead to a change in background for the new beneficial mutation. Similar formulae can be developed for the other $P_{mk}[t]$.

By assuming that selection is weak and that the fixation probabilities P_{mk} are $O(s)$, we can rewrite (1) as

$$\begin{aligned}\Delta P_{11} = P_{11}[t + 1] - P_{11}[t] \approx &-(s + Sq_k)P_{11} \\ &+ P_{11}^2/2 + (P_{11} - P_{10})(1 - P_{11} + s + q_kS) \\ &\times ((1 - q_mr_1)q_k^A(r_2 + \delta r_2(q_m - p_m)) + \delta r_2\, p_mq_k^A) \\ &+ (P_{11} - P_{01})(1 - P_{11} + s + q_kS)q_mr_1(1 - q_k^A(r_2 \\ &+ \delta r_2(q_m - p_m))) + (P_{11} - P_{00})(1 - P_{11} + s + q_kS) \\ &\times q_mr_1q_k^A(r_2 + \delta r_2(q_m - p_m)) + \tfrac{1}{2}[(P_{11} - P_{01})q_mr_1 \\ &\times (1 - q_k^Ar_2) + (P_{11} - P_{10})q_k^Ar_2(1 - q_mr_1) \\ &+ (P_{11} - P_{00})q_mq_k^Ar_1r_2]^2 + o(s^2) + o(\delta r_2). \quad (3)\end{aligned}$$

When selection is weak, the change in P_{mk} over time will be small and we can make the continuous time approximation, $dP_{mk}/dt \approx P_{mk}[t + 1] - P_{mk}[t]$. Since the modifier is assumed to be weak, it will only cause slight changes in the fixation probabilities. We can then write $P_{11} = P_1 + q_m\,\delta P_1$, $P_{10} = P_0 + q_m\,\delta P_0$, $P_{01} = P_1 - p_m\,\delta P_1$, and $P_{00} = P_0 - p_m\,\delta P_0$. P_k represents the fixation probability of j_1 when linked with allele k averaged over the two backgrounds at the modifier locus ($P_k = p_mP_{1k} + q_mP_{0k}$), whereas δP_k represents the difference in fixation probability of j_1 between the two backgrounds at the modifier locus ($\delta P_k = P_{1k} - P_{0k}$).

To analyze these equations further, we follow the procedure described in BARTON (1995b). Whether or not the new beneficial mutation (j_1) will become established is determined largely while it is still rare, during which time it will not greatly perturb the dynamics at locus K. Assuming that the beneficial allele k_1 is already sufficiently frequent that its dynamics are deterministic and assuming that selection is relatively weak at locus K, the frequency of k_1 can be described by the logistic equation,

$$p_k = \frac{e^{St}}{1 + e^{St}} \quad q_k = \frac{1}{1 + e^{St}}, \quad (4)$$

where time is measured relative to the mid-point of the sweep of allele k_1 ($p_k = {}^1/_2$ at $t = 0$). The probability of fixation of j_1 can then be determined using these equations for p_k and q_k.

To simplify the equations, we employ the following scalings (BARTON 1995b):

$$\Pi = \frac{(p_kP_1 + q_kP_0)}{2s} \quad \Delta = \frac{(P_1 - P_0)}{2s}. \quad (5)$$

Π measures the probability of fixation averaged across all backgrounds relative to the expected fixation probability from one locus theory ($2s$). Δ measures the difference in fixation probability between the two genetic backgrounds (with k_1 and with k_0) also relative to $2s$. We will measure variation in these quantities due to the modifier locus by

$$\delta\Pi = \frac{S(p_k\,\delta P_1 + q_k\,\delta P_0)}{2s\,\delta r_2} \quad \delta\Delta = \frac{S(\delta P_1 - \delta P_0)}{2s\,\delta r_2}. \quad (6)$$

$\delta\Pi$ measures the amount by which the probability of fixation is increased by the modifier, while $\delta\Delta$ measures changes in Δ caused by the modifier. Finally, time and the other variables are scaled as

$$T = St \quad \rho_1 = \frac{r_1}{S} \quad \rho_2 = \frac{r_2}{S} \quad \chi = \frac{r_1r_2}{S} \quad \theta = \frac{s}{S}.$$

Using these scalings and the equations for dP_{mk}/dt, we can obtain equations describing how Π, Δ, $\delta\Pi$, $\delta\Delta$ change over time. Assuming that recombination is weak (specifically that r_2 is not much greater than S) and retaining only leading order terms,

$$\frac{d\Pi}{dT} = -\theta\Pi(1 - \Pi) + \theta p_kq_k\Delta^2 + O(s) \quad (7)$$

$$\frac{d\Delta}{dT} = \Delta(\rho_2 + (2\Pi - 1)\theta + (p_k - q_k)(1 - \theta\Delta)) - \Pi + O(s). \quad (8)$$

$$\frac{d\delta\Pi}{dT} = (\rho_1 + \theta(2\Pi - 1))\delta\Pi + 2\theta p_kq_k\Delta\,\delta\Delta + O(s) \quad (9)$$

$$\frac{d\delta\Delta}{dT} = (\rho_1 + \rho_2 - \chi + \theta(2\Pi - 1))\delta\Delta + (1 - 2\theta\Delta)((p_k - q_k)\delta\Delta - \delta\Pi) + \Delta + O(s). \quad (10)$$

Equations 7 and 8 are the same as those in the absence of a modifier (Equations 6a and 6b in BARTON 1995b). For loose recombination (when ρ_2 is large), some of the $O(s)$ terms can no longer be neglected. More exact equations that also apply for loose recombination are provided in APPENDIX B.

The probability of fixation of the beneficial j_1 allele when it arises with m_1 is related to the above quantities by the following formula,

$$F_1 = 2s\left(\Pi + \frac{q_m\, \delta r_2\, \delta\Pi}{S}\right), \quad (11)$$

which is an average fixation probability with respect to the alleles at locus K. Similarly, if j_1 arises with allele m_0, its probability of fixation is

$$F_0 = 2s\left(\Pi - \frac{p_m\, \delta r_2\, \delta\Pi}{S}\right). \quad (12)$$

Either the exact equations (1) or the approximate equations (7)–(10) may be solved numerically, *e.g.*, using the Runge-Kutta method employed by NDSolve of Mathematica (WOLFRAM RESEARCH 1993) (further details are available in BARTON 1995b). These equations describe the impact of directional selection in the background on the probability of fixation of a new beneficial mutation as a function of time. They also describe the effect of a modifier on the average probability of fixation. In Figure 1, Π and $\delta\Pi$ are shown as a function of the time T when the j_1 mutation appears [recall that since $T = St$, $p_k = e^T/(1 + e^T)$ and thus T can be used to measure the frequency of k_1 when j_1 appears]. In general, the probability of fixation of j_1 remains near $2s$ if the mutation arises while K is still uncommon ($T \to -\infty$) or if it arises when K is near fixation ($T \gg 0$). When there is substantial genetic variation at the K locus, however, the probability of fixation can be reduced dramatically, especially if selection is weak at locus J relative to locus K ($s/S = \theta \ll 1$). We turn now to approximate solutions that can be used when linkage is either tight or loose. Afterwards, we will use these approximations to calculate the expected change in frequency of a modifier allele.

A weakly selected mutation under tight linkage: In this section, we focus on the case in which the modifier will have the most dramatic influence on the probability of fixation of j_1: small θ ($=s/S$) and small ρ_i ($=r_i/S$). The details are presented in APPENDIX A, where we show that

$$\delta\Pi \approx \frac{\partial\Pi}{\partial\rho_2} e^{\rho_1 T} = \frac{e^{(\rho_1+\theta)T}\theta^{\rho_2} \log\left(\frac{1}{\theta}\right)}{(1 - \theta^{\rho_2} + e^{\theta T}\theta^{\rho_2})^2}, \quad (13)$$

which employs Equation A1 of BARTON (1995b):

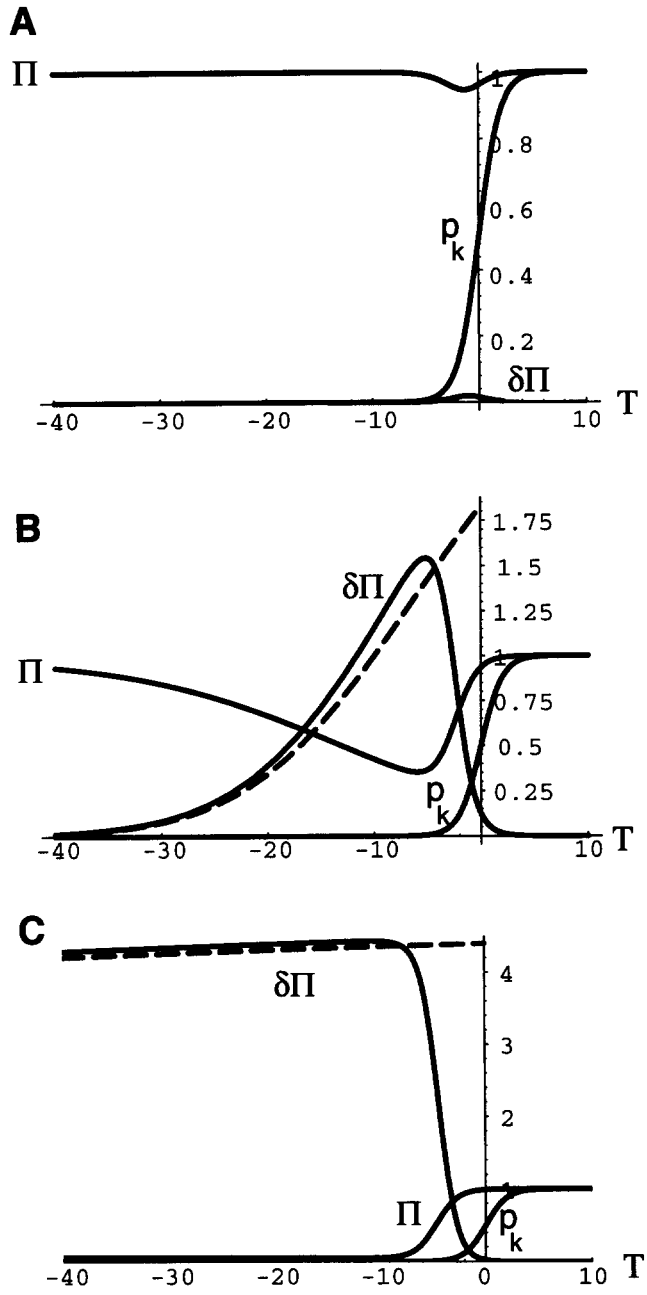

FIGURE 1.—Decreasing the size of ρ_1, ρ_2, θ increases the importance of background selection. The solid curves show a numerical evaluation of (7)–(10). Π measures the amount by which fixation is reduced from its expectation. $\delta\Pi$ measures the increase in fixation probability due to a modifier scaled by $2s\,\delta r_2/S$. p_k is the frequency of the favorable allele at locus K. All are graphed as a function of T, the time at which the new beneficial allele appears at locus J. The dashed curves give an approximate solution for $\delta\Pi$ from (13). (A) $\rho_1 = \rho_2 = \theta = 1$. (B) $\rho_1 = \rho_2 = \theta = 0.1$. (C) $\rho_1 = \rho_2 = \theta = 0.01$.

$$\Pi \approx \frac{1 - \theta^{\rho_2}}{1 - \theta^{\rho_2} + \theta^{\rho_2} e^{\theta T}}, \quad (14)$$

which is accurate for small θ and ρ_2, and for $T < 0$. As T approaches zero, approximation (13) becomes invalid. In this case, $\delta\Pi$ returns rapidly to zero. Figures 1 and 2 illustrate that (13) does provide a good approxi-

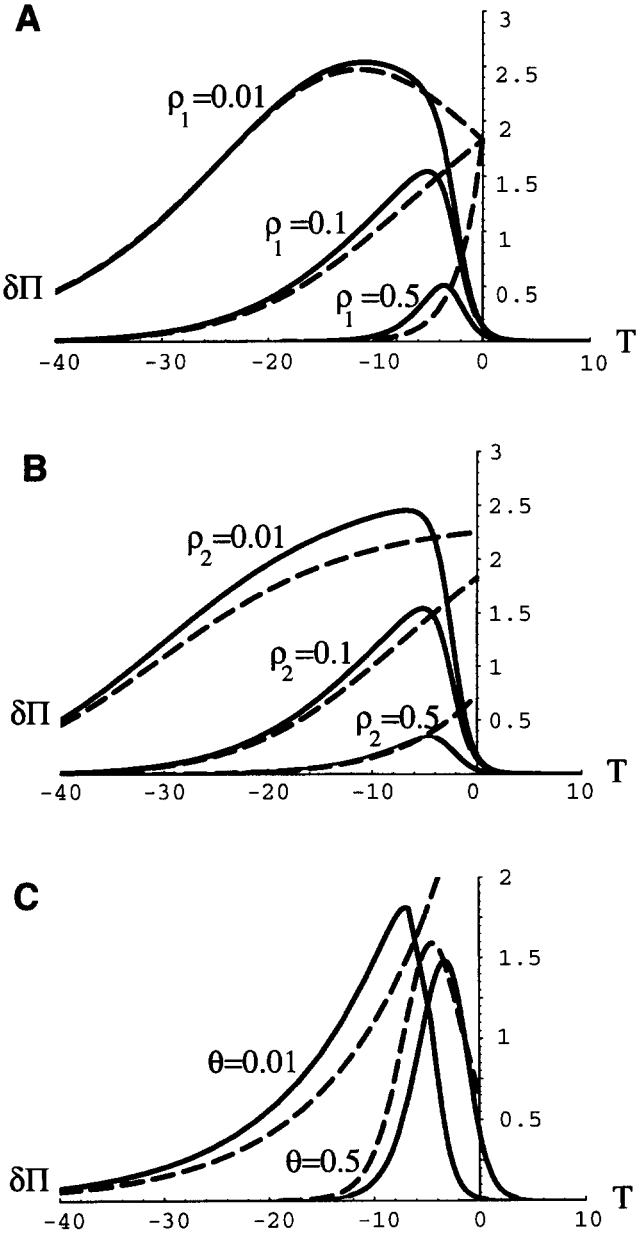

FIGURE 2.—Decreasing either recombination among loci or the selection coefficient of a new mutation increases the effect of a modifier on the probability of fixation of the mutation, shown here as increases in $\delta\Pi$. Solid curves are from a numerical evaluation of (7)–(10). Dashed curves are the approximate solutions for $\delta\Pi$ from (13). (A) $\rho_2 = \theta = 0.1$. (B) $\rho_1 = \theta = 0.1$. (C) $\rho_1 = \rho_2 = 0.1$.

mation to $\delta\Pi$ except near $T = 0$. As expected, as T goes to negative infinity, the effect of the modifier disappears ($\delta\Pi \to 0$). To help interpret these results, note that the difference between the fixation probability of j_1 when linked with m_1 and when linked with m_0 is $F_1 - F_0 = 2s(\delta r_2/S)\delta\Pi$. Since (13) for $\delta\Pi$ is always positive, an initial association with a modifier allele that increases recombination will always increase the average fixation probability of a new beneficial mutation. The amount by which it does so is proportional to δr_2, such that the stronger the modifier, the more likely it is for a beneficial mutation to fix within the population.

As a numerical example, consider $r_1 = r_2 = s = 0.01$ and $S = 0.1$ (that is, $\theta = \rho_1 = \rho_2 = 0.1$ as in Figure 1B) with a modifier frequency of $p_m = 0.5$. If j_1 appears when $p_k = 4.5 \times 10^{-5}$ ($T = -10$), Π is ~0.4 and $\delta\Pi$ is ~1. Therefore, selection at locus K reduces the fixation probability of j_1 from the expectation of $2s = 0.02$ to ~0.008. The fixation probability of j_1 when it is initially associated with m_1 is, however, higher than when it is initially associated with m_0 by an amount $F_1/F_0 \approx 1 + (\delta r_2 \delta\Pi/\Pi S) = 1 + 25\,\delta r_2$. The impact of a modifier is even more pronounced when linkage is tighter. If in the above example r_1 and r_2 are reduced to 0.001, the average fixation probability of j_1 is only about 0.001. Yet, the fixation probability of j_1 when initially linked with m_1 is about $1 + 2000\,\delta r_2$ times higher than when initially linked with m_0. Clearly, a modifier allele that increases recombination can substantially increase the probability of fixation of a beneficial mutation. In general, the impact of a modifier on fixation probability will be greatest when linkage is tight.

A measure that will become important in determining the expected change in frequency of a modifier is the net effect of the modifier on the probability of fixation found by integrating $\delta\Pi$ over all possible times of origin of new beneficial mutations (see APPENDIX A). There we show that

$$\int_{-\infty}^{\infty} \delta\Pi dT \approx \frac{\log[\theta]\left(\rho_1 F\left[1, 1+\frac{\rho_1}{\theta}, 2+\frac{\rho_1}{\theta}, -\frac{\theta^{\rho_2}}{1-\theta^{\rho_2}}\right] - (\rho_1+\theta)(1-\theta^{\rho_2})\right)}{(\rho_1+\theta)\theta^{(1-\rho_2)}(1-\theta^{\rho_2})^2}, \quad (15)$$

where F is the hypergeometric function. Equation 15 is compared to numerical solutions of (1) in Table 3. As a rough approximation, the net effect is on the order of $1/(\rho_2\theta) = S^2/(sr_2)$ when recombination is rare within the population ($r_1, r_2 \ll s$).

Loose linkage: Although the strongest effect of a modifier on fixation probability is seen with tight linkage, it is important to assess the impact that a modifier may have in a genome that already has substantial recombination. Although we were unable to solve for $\delta\Pi$ directly, it was possible to get an estimate for the net effect of the modifier as shown in APPENDIX B. With loose linkage ($r_1, r_2 \gg S$), the net effect is given by (76), which can be written in terms of the unscaled variables as

$$\int_{-\infty}^{\infty} \delta\Pi dT \approx \frac{(r_1 r_2 + 2s + r_1 s)S^3}{(r_2+s)^2(r_1+r_2-r_1r_2+s)(r_1+s-r_1s)}. \quad (16)$$

Equation 16 is also compared to numerical solutions of

TABLE 3

Net effect of a modifier on the fixation probability

r_1	r_2	s	Exact[a]	Small r approx.[b]	Large r approx.[c]	Large r approx.[d]
10^{-4}	10^{-4}	10^{-4}	1031.18	1034.81	NA	83348.6
10^{-4}	10^{-4}	10^{-3}	682.71	680.51	NA	1252.38
10^{-4}	10^{-4}	10^{-2}	37.91	NA	3.72	1.90
10^{-4}	10^{-4}	10^{-1}	0.0022	NA	0.0020	0.0020
10^{-4}	10^{-2}	10^{-4}	1.77	2.32	1.89	0.97
10^{-4}	10^{-2}	10^{-3}	1.55	2.31	2.37	1.35
10^{-4}	10^{-2}	10^{-2}	0.24	NA	0.29	0.25
10^{-4}	10^{-2}	10^{-1}	0.0016	NA	0.0015	0.0015
10^{-2}	10^{-4}	10^{-4}	1.38	4.44	NA	49.03
10^{-2}	10^{-4}	10^{-3}	1.37	2.49	NA	13.63
10^{-2}	10^{-4}	10^{-2}	1.47	NA	0.82	0.49
10^{-2}	10^{-4}	10^{-1}	0.0018	NA	0.0017	0.0017
10^{-2}	$1/2$	10^{-4}	$0.57\ 10^{-5}$	NA	$0.41\ 10^{-5}$	$0.41\ 10^{-5}$
10^{-2}	$1/2$	10^{-3}	$0.69\ 10^{-5}$	NA	$0.50\ 10^{-5}$	$0.50\ 10^{-5}$
10^{-2}	$1/2$	10^{-2}	$1.21\ 10^{-5}$	NA	$0.94\ 10^{-5}$	$0.94\ 10^{-5}$
10^{-2}	$1/2$	10^{-1}	$1.21\ 10^{-5}$	NA	$0.87\ 10^{-5}$	$0.87\ 10^{-5}$
$1/2$	10^{-2}	10^{-4}	$2.03\ 10^{-4}$	NA	$2.07\ 10^{-4}$	$2.04\ 10^{-4}$
$1/2$	10^{-2}	10^{-3}	$*2.21\ 10^{-4}$	NA	$2.67\ 10^{-4}$	$2.45\ 10^{-4}$
$1/2$	10^{-2}	10^{-2}	$*2.70\ 10^{-4}$	NA	$3.06\ 10^{-4}$	$2.88\ 10^{-4}$
$1/2$	10^{-2}	10^{-1}	$0.64\ 10^{-4}$	NA	$0.63\ 10^{-4}$	$0.63\ 10^{-4}$
$1/2$	$1/2$	10^{-4}	$2.67\ 10^{-6}$	NA	$2.67\ 10^{-6}$	$2.67\ 10^{-6}$
$1/2$	$1/2$	10^{-3}	$2.68\ 10^{-6}$	NA	$2.68\ 10^{-6}$	$2.68\ 10^{-6}$
$1/2$	$1/2$	10^{-2}	$2.79\ 10^{-6}$	NA	$2.75\ 10^{-6}$	$2.75\ 10^{-6}$
$1/2$	$1/2$	10^{-1}	$3.17\ 10^{-6}$	NA	$2.97\ 10^{-6}$	$2.97\ 10^{-6}$

Estimates are given for $\int_{-\infty}^{\infty} \delta\Pi dT$, concentrating on parameters for which the approximations might break down (θ, ρ_1, or ρ_2 near 1). S is set to 0.01 and $\delta r_2 = 10^{-6}$ throughout.

[a] Exact numerical integration of $\int_{-\infty}^{\infty} \delta\Pi dT$ using formula (1) with definitions (5) and (6). Difficulties with the numerical solutions were experienced in cases marked by an asterisk (these values may be inaccurate).

[b] Tight linkage approximation using (58).

[c] Loose linkage approximation using (75).

[d] Loose linkage approximation using the much simpler (76).

(1) in Table 3. The net effect of a modifier on the probability of fixation decreases rapidly as recombination rates increase. As recombination rates approach $1/2$, the net effect of a modifier on the probability of fixation of j_1 declines to $\approx 8S^3/3$, at which point the modifier has a negligible effect on the fate of a new mutation.

Dynamics at the modifier locus: As beneficial mutations sweep through a population, they increase the frequency of linked alleles in a process known as "hitchhiking" (MAYNARD-SMITH and HAIGH 1974). By making it more likely that a favorable mutation will rise to fixation, a modifier allele that increases recombination rates between selected loci is more likely to gain the benefit of a hitch than is a modifier allele that decreases recombination. In this section, we describe the expected change in frequency at a modifier locus by taking into account both the fixation probability of beneficial mutations on different modifier backgrounds and also the subsequent change in frequency at the modifier locus due to hitchhiking. An examination of this process addresses our central question: to what extent will higher rates of recombination evolve because recombination increases the fixation probability of adaptive mutations?

We assume that the modifier allele is selectively neutral except for its effects on the fixation probability of new mutations; that is, we focus on allele frequency changes that occur at the modifier locus due to changes in the fixation probability of adaptive mutations and ignore other possible selective forces acting upon the modifier locus [such as selection that arises in the presence of epistatic interactions among selected loci (BARTON 1995a)]. The first step is to determine how much a modifier allele is expected to change in frequency given that it is initially linked to a beneficial allele that is destined to fix. In APPENDIX C, we develop an approximate solution to the expected change in frequency of a neutral allele due to hitchhiking with a beneficial allele. Although the effects of hitchhiking on neutral variability have been examined before, previous studies have focused on the expected change in heterozygosity

(STEPHAN *et al.* 1992) or have used a purely deterministic approach (MAYNARD-SMITH and HAIGH 1974). The approach we take combines results for the stochastic dynamics that occur while the new mutation is rare with results from a deterministic analysis once the allele is common. The initial stochastic effects are important because those mutations that are successful tend to rise rapidly in frequency during the first few generations and this acceleration has a strong effect on linked neutral loci. As shown in APPENDIX C, the expected change in allele frequency of the neutral allele m_1 when it is initially linked to a new beneficial mutation (j_1) is

$$\Delta_1 p_m \approx q_m (4Ns)^{-r_1/s} \Gamma\left(1 + \frac{r_1}{s}\right)^2 \Gamma\left(1 - \frac{r_1}{s}\right) \quad \text{for tight linkage and} \quad (17)$$

$$\approx q_m \frac{1}{4N(r_1 - s)} \quad \text{for loose linkage,} \quad (18)$$

where the change in allele frequency is measured over the entire period during which j_1 increases in the population and is conditional on the fixation of j_1. This must be multiplied by the probability of fixation of j_1 when initially linked with m_1 (given by F_1, Equation 11) to get the expected change in the frequency of m_1 averaged over both processes that do lead to fixation and those that do not. A similar equation for the change in m_1 when j_1 initially appears with m_0 ($\Delta_0 p_m$) is presented in APPENDIX C.

Overall, the expected change in frequency of the modifier allele m_1 is equal to

$$\Delta p_m = p_m \Delta_1 p_m F_1 + q_m \Delta_0 p_m F_0$$

$$\approx \frac{2s\,\delta r_2\,\delta\Pi}{S} p_m q_m (4Ns)^{-r_1/s} \Gamma\left(1 + \frac{r_1}{s}\right)^2 \Gamma\left(1 - \frac{r_1}{s}\right) \quad \text{for tight linkage and} \quad (19)$$

$$\approx \frac{2s\,\delta r_2\,\delta\Pi}{S} p_m q_m \frac{1}{4N(r_1 - s)} \quad \text{for loose linkage,} \quad (20)$$

where the expectation is taken over both cases where j_1 initially arises with m_1 and cases where j_1 arises with m_0 and takes into account the probability of fixation of j_1 in both these cases.

For any particular set of parameters, $\delta\Pi$ in the above equations may be found numerically using (7) – (10). With tight linkage and weak selection acting on the new mutation, we may use (13) to find that the expected change in frequency at the modifier locus due to hitchhiking with the new mutation is

$$\Delta p_m \approx 2\,\delta r_2\,\theta \left(\frac{e^{(\rho_1+\theta)T}\theta^{\rho_2} \log\left(\frac{1}{\theta}\right)}{(1 - \theta^{\rho_2} + e^{\theta T}\theta^{\rho_2})^2} \right) \times \left(p_m q_m (4Ns)^{-r_1/s} \Gamma\left(1 + \frac{r_1}{s}\right)^2 \Gamma\left(1 - \frac{r_1}{s}\right) \right). \quad (21)$$

The expected change at the modifier locus decreases rapidly with increasing recombination as shown in Figure 3. The largest expected change in frequency of the modifier allele occurs as recombination rates go to zero ($\rho_1, \rho_2 \to 0$), in which case

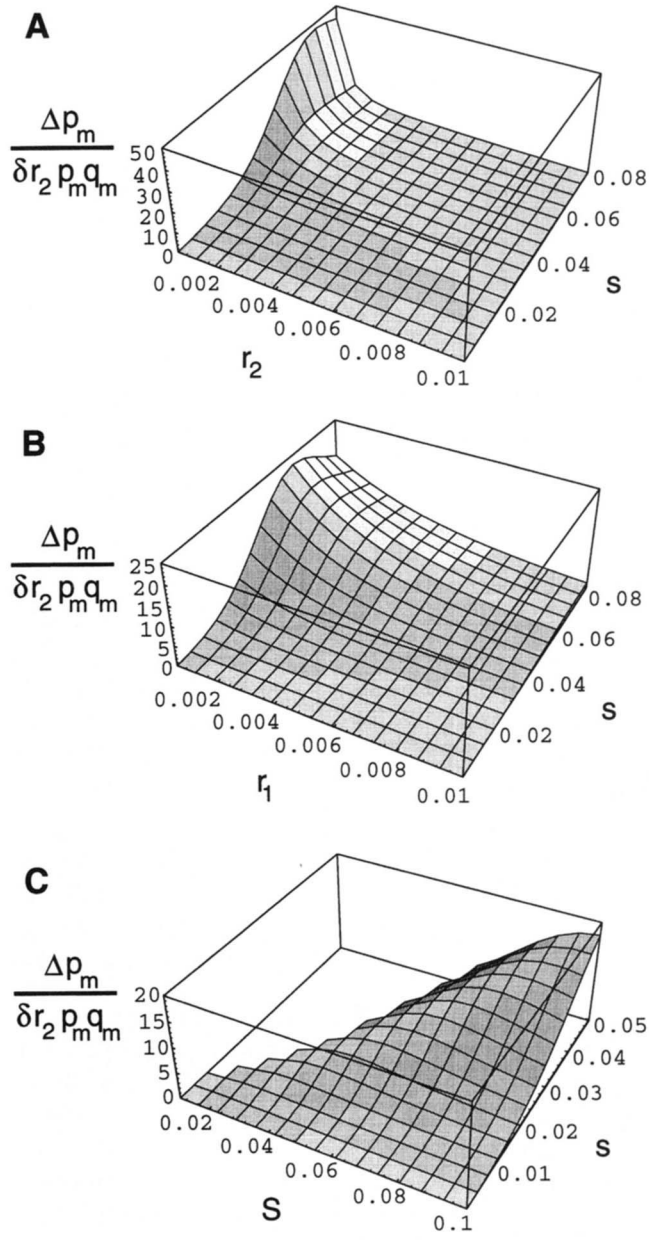

FIGURE 3.—The expected change in frequency of a modifier allele decreases rapidly with increasing recombination. The function (21) is used to show how the expected change in frequency at the modifier locus (Δp_m) varies as a function of r_1, r_2, S, and s using $T = -10$ and $N = 10^6$. (A) $r_1 = 0.001$, $S = 0.1$. (B) $r_2 = 0.001$, $S = 0.1$. (C) $r_1 = 0.001$, $r_2 = 0.001$.

$$\frac{\Delta p_m}{p_m q_m} \approx \delta r_2 \frac{2\theta \log\left[\frac{1}{\theta}\right]}{e^{\theta T}} = \delta r_2 \frac{2\theta \log\left[\frac{1}{\theta}\right]}{[p_k/q_k]^{\theta}}. \quad (22)$$

This is the expected change at the modifier locus when j_1 appears at time T through hitchhiking whether or not j_1 fixes. This approximation assumes that $r_1, r_2 \ll s \ll S$ and that $p_k < {}^1/_2$. The modifier will change most in frequency when recombination rates are low and when the mutation appears while k_1 is still rare.

Genome wide mutations: We now generalize the process considered above by assuming that beneficial mutations arise anywhere along a chromosome throughout the population at a certain rate, λ. Although this rate may be high, technically we assume that the rate of successful mutations (ones that become established) is low. That is, we assume that there is rarely more than one segregating locus in the background when new mutations appear within the population. This restriction may be unnecessary since it has been shown that, with multiple substitutions, the fixation probability of a particular mutation declines linearly with λ up to a threshold at which point fixation is very unlikely (BARTON 1994, 1995b). Therefore it is plausible that the effect of a modifier on the probability of fixation of a particular mutation increases linearly with λ even when multiple substitutions are in progress, at least up to a point. This claim needs to be verified, however.

The expected change in frequency at the modifier locus depends on how frequently there is directional selection at other loci when a new mutation arises. The rate at which successful mutations appear in the population is approximately $\lambda E[2S\Pi^*]$, where Π^* is the average extent to which the probability of fixation is decreased from $2S$ by multiple substitutions and is given by

$$\Pi^* = 1 - \lambda \int_{-\infty}^{\infty} (1 - \Pi)\, dt = 1 - \lambda/S \int_{-\infty}^{\infty} (1 - \Pi)\, dT, \quad (23)$$

assuming that $\lambda < S\int_{-\infty}^{\infty}(1-\Pi)\,dT$ (BARTON 1995b). Any particular successful mutation (with advantage S) will then segregate within the population for a period of time, say between t_0 and t_1. Therefore the fraction of time during which beneficial alleles are segregating within the population is approximately $\lambda E[2S\Pi^*(t_1 - t_0)]$. The rate at which new mutations occur within the population during periods when previous beneficial mutations are still segregating is thus $\lambda^2 E[2S\Pi^*(t_1 - t_0)]$. Most of these mutations will fail to become established within the population, but each of them will contribute to the expected change at the modifier locus. For each new mutation that arises between t_0 and t_1, the expected change in frequency of the modifier allele, m_1, will be

$$E\left[\int_{t_0}^{t_1} \frac{\Delta p_m}{t_1 - t_0}\, dt\right],$$

assuming that mutations occur uniformly over the interval (t_0, t_1). Overall, the expected change at the modifier locus per generation is

$$\overline{\Delta p_m} = E\left[\lambda^2 (2S\Pi^*(t_1 - t_0)) \int_{t_0}^{t_1} \frac{\Delta p_m}{t_1 - t_0}\, dt\right] = E\left[\lambda^2\, 2S\Pi^* \int_{t_0}^{t_1} \Delta p_m\, dt\right].$$

Scaling time according to $T = St$ with $T = 0$ at the midpoint of the sweep, we have that

$$\overline{\Delta p_m} = E\left[\lambda^2 2S\Pi^* \int_{T_0}^{T_1} \frac{\Delta p_m}{S}\, dT\right] \approx E\left[\lambda^2\, 2\Pi^* \int_{-\infty}^{\infty} \Delta p_m\, dT\right], \quad (24)$$

which takes into account the fact that Δp_m is extremely small when $T \ll 0$ and $T \gg 0$ and so there is a negligible effect of changing the limits of integration.

When linkage throughout the genome is tight, we can replace Δp_m in the above equation with (19) to obtain

$$\overline{\Delta p_m} \approx E\left[4\Pi^*\lambda^2\delta r_2\, \theta\left(p_m q_m (4Ns)^{-r_1/s} \Gamma\left(1 - \frac{r_1}{s}\right)\Gamma\left(1 + \frac{r_1}{s}\right)^2\right) \int_{-\infty}^{\infty} \delta\Pi\, dT\right]. \quad (25)$$

This equation can be solved numerically or by using approximation (15) for small θ:

$$\overline{\Delta p_m} \approx \left[E\ 4\Pi^*\lambda^2\delta r_2\theta\left(p_m q_m (4Ns)^{-r_1/s} \Gamma\left(1 - \frac{r_1}{s}\right)\Gamma\left(1 + \frac{r_1}{s}\right)^2\right) \times \left(\frac{\log[\theta]\left(\rho_1 F\left[1, 1 + \frac{\rho_1}{\theta}, 2 + \frac{\rho_1}{\theta}, -\frac{\theta^{\rho_2}}{1 - \theta^{\rho_2}}\right] - (\rho_1 + \theta)(1 - \theta^{\rho_2})\right)}{(\rho_1 + \theta)\theta^{(1-\rho_2)}(1 - \theta^{\rho_2})^2}\right)\right]. \quad (26)$$

An upper bound for the expected change in modifier frequency can be found by assuming that the recombination rates are extremely small. Then, to leading order

$$\overline{\Delta p_m} \approx E\left[4\Pi^*\lambda^2\,\delta r_2\,p_m q_m\frac{S}{r_2}\right]$$
$$\approx E\left[4\left(1-\frac{\lambda}{S}\frac{-\log\,[1-\theta^{\rho_2}]}{\theta}\right)\lambda^2\,\delta r_2\,p_m q_m\frac{S}{r_2}\right] .$$

For low rates of beneficial mutation, this implies that indirect selection is occurring at the modifier locus with a selection coefficient *per generation* of

$$\overline{s_m} = \frac{\overline{\Delta p_m}}{p_m q_m} \approx 4\,\bar{S}\,\lambda^2 E\left[\frac{\delta r_2}{r_2}\right], \qquad (27)$$

where $\bar{S}$ is the average selection coefficient of new mutations that is assumed to be independent of the genetic map. As one might expect, if the rate of mutation to beneficial alleles is extremely small per generation, then interference among positively selected loci is negligible and there will be no selection acting at the modifier locus. However, if it is reasonably common for new beneficial mutations to occur while a previous mutation is still segregating and if recombination rates are low, the modifier can have a large impact on the rate of fixation of new beneficial mutations and experience substantial positive selection.

When linkage throughout the genome is loose, we can replace Δp_m in (24) using (20) to obtain

$$\overline{\Delta p_m} \approx E\left[2\Pi^*\lambda^2\,\delta r_2\,p_m q_m\frac{2s}{S}\frac{1}{4N(r_1-s)}\int_{-\infty}^{\infty}\delta\Pi\,dT\right], \qquad (28)$$

which becomes

$$\overline{\Delta p_m} \approx E\left[2\Pi^*\lambda^2\,\delta r_2\,p_m q_m\frac{2s}{S}\frac{1}{4N(r_1-s)}\times\frac{(r_1r_2+2s+r_1s)\,S^3}{(r_2+s)^2(r_1+r_2-r_1r_2+s)\,(r_1+s-r_1s)}\right], \qquad (29)$$

using (16) for the net effect with loose linkage. The effect will be minimal when nearly all loci are unlinked (r_1, $r_2 \approx {}^1/_2$), in which case the expected strength of selection acting on the modifier per generation becomes

$$\overline{s_m} = \frac{\overline{\Delta p_m}}{p_m q_m} \approx E\left[\lambda^2\,\delta r_2\frac{16sS^2}{3N}\right],$$
$$\approx \frac{16}{3N}\lambda^2 E[sS^2]\,E[\delta r_2],$$

which is clearly small.

When recombination rates are determined by the position of genes along a chromosome of map length R, we must integrate the change in modifier frequency over all possible values of r_1 and r_2 to assess the change expected on a random genetic map. To simplify matters, we assume that the chromosome is fairly long (or circular) so that we may ignore edge effects. In this case, choosing the recombination rate between loci M and J places no constraints on the recombination rate between J and K. Furthermore, to simplify the calculations, we assume that r_1 and r_2 will be uniformly distributed between some small value (say r_{small}) and ${}^1/_2$ with probability $1/(2R)$ and will be ${}^1/_2$ otherwise. Making these simplifying assumptions (which should not qualitatively alter our estimates), the expected change in frequency of the modifier allele, m_1, is given by

$$\overline{\Delta p_m} = \left(\frac{1}{2R}\right)^2\int_{r_{small}}^{1/2}\int_{r_{small}}^{1/2}\Delta p_m\,dr_1\,dr_2$$
$$+\left(1-\frac{1}{2R}\right)\frac{1}{2R}\left(\int_{r_{small}}^{1/2}(\Delta p_m|\,r_2={}^1/_2)\,dr_1 + \int_{r_{small}}^{1/2}(\Delta p_m|\,r_1={}^1/_2)\,dr_2\right)$$
$$+\left(1-\frac{1}{2R}\right)^2\times(\Delta p_m|\,r_1=r_2={}^1/_2). \qquad (30)$$

To evaluate these integrals, we note that (29) is approximately equal to

$$\overline{\Delta p_m} \approx E\left[2\Pi^*\lambda^2\,\delta r_2\,p_m q_m\frac{2s}{S}\frac{1}{4Nr_1}\times\frac{(r_1r_2+2s+r_1s)\,S^3}{(r_2+s)^2(r_1+r_2-r_1r_2+s)\,r_1}\right]. \qquad (31)$$

This equation is not only an excellent approximation for the change in allele frequency with loose linkage ($r_1 > s$), but it also provides an estimate for the change in allele frequency when recombination is tight. For tight linkage, (31) is off, often by orders of magnitude (see Figure 10), but tends to underestimate Δp_m estimated using (27) except when recombination rates are extremely low. Therefore, for our purposes, we can use (31) to obtain a conservative estimate for the expected change at a modifier locus with genes placed at random along a chromosome.

Keeping only leading order terms in r_{small}, s, and S for each of the four integrals in (30) with (31), we find that approximately

$$\overline{\Delta p_m} \approx 2\Pi^*\lambda^2\frac{\delta r_2\,p_m q_m}{2N}\left[\left(\frac{1}{2R}\right)^2\frac{S^2}{r_{small}} + \left(1-\frac{1}{2R}\right)\times\frac{1}{2R}\left(\frac{16s^2S^2}{r_{small}}+4sS^2(3-\log[s])\right) + \left(1-\frac{1}{2R}\right)^2\frac{16}{3}sS^2\right], \qquad (32)$$

where we have dropped the $E[\ ]$ notation for clarity. Except when the map length is extremely large, (32)

will be dominated by the case when all loci are linked (*i.e.*, the first term).

To evaluate this result, we draw rough estimates of the parameters from data on humans and Drosophila (GRIFFITHS *et al.* 1996, p. 136, 192, 494). There are $\sim 3.3 \times 10^9$ bp in the human genome, corresponding to about 33 morgans per genome and 1.4 morgans per chromosome. Similarly, chromosomes *2*, *3*, and *X* of Drosophila average 0.9 morgans in length. We thus take the map length of an average chromosome, R, to be about 1 morgan. To estimate r_{small}, we note that with $\sim 10^5$ coding regions in the human genome, the recombination distance between adjacent loci is an average of 3.3×10^{-4} morgans. Drosophila, with an estimated 10^4 genes, provide a similar estimate of $r_{small} \approx 2.7 \times 10^{-4}$ morgans. These estimates are cautious since they ignore the possibility of multiple beneficial mutations within the same gene. Data from Drosophila indicate that, for deleterious mutations, average selection coefficients are on the order of 0.01 (MUKAI *et al.* 1972), which we take as an estimate of s and S although little is known about the selective advantage of typical beneficial mutations (ORR and COYNE 1992). Taken together with (32), an extremely rough estimate for the average strength of selection favoring recombination at a modifier locus due to changes in the fixation probability of mutations on the same chromosome is

$$s_m \approx \frac{\overline{\Delta p_m}}{p_m q_m} \approx \frac{\Pi^* \lambda^2 \delta r_2}{12N}. \qquad (33)$$

λ, the rate of appearance of beneficial mutations on the chromosome, is unknown. Based on rates of amino acid substitutions in hemoglobin genes, MAYNARD SMITH and HAIGH (1974) estimated the rate of substitution of beneficial mutations in vertebrates to be ~ 0.06 per genome per generation. If approximately $2s$ of all the beneficial mutations that occur lead to a substitution and if s is about 0.01, the total rate of appearance of beneficial mutations within a population would be about six per genome per generation (which, for humans, would translate to about $1/4$ per chromosome per generation). Under these circumstances, selection on the modifier is extremely weak ($\ll 1/N$) and would be overwhelmed by drift. For selection to dominate over drift, the rate of appearance of beneficial mutations, λ, must be greater than $\sqrt{12/\delta r_2}$, assuming that Π^* is near one. This condition may be too pessimistic, since we have underestimated the importance of tightly linked loci by using (51) rather than (27) and by using an r_{small} of 3×10^{-4}. Furthermore, since there are thought to be more than 20 loci modifying recombination rates in Drosophila (BAKER *et al.* 1976), the total response to selection for increased recombination rates may be substantial even when the selection coefficient per modifier locus is weak.

Even if beneficial mutations do not always occur at a high rate, there may be occasional periods of intense directional selection during which time the rate of beneficial mutation is increased. Such periods are expected to occur under the punctuated equilibrium model of evolution during bouts of rapid evolutionary change. Intense directional selection is also typical of many selection experiments. If, under conditions of strong directional selection, λ and S both increase by an order of magnitude (that is, more loci are under selection and selection is stronger), selection on a modifier locus would increase by four orders of magnitude and would much more strongly favor the increase of recombination. Of course, in selection experiments, populations tend to be small, so that λ, the rate of appearance of new beneficial mutations in the population, is likely to be very much decreased during the course of the experiment. Nevertheless, alleles present in the stock population that become beneficial under the selective conditions of the experiment would act like newly arisen mutations (as long as they are not sufficiently frequent to be assured of fixation) and would contribute to λ. In the experiment of KOROL and ILIADI (1994), using the parameter values of 2.5 for λ, 0.1 for S, and 28.8 for N_e (as in their experiment), selection on each modifier locus would have been $s_m \approx 1.8\,\delta r_2$. Using (11.8) and (11.3) of FALCONER (1989), the expected response to selection under these conditions would be $R = 1.8V_A$, where V_A is the additive genetic variance for recombination. Although we do not know the additive genetic variance for recombination in this population, it is certainly plausible that recombination rates could have risen by 33% in 50 generations in this selection experiment on geotaxis as a consequence of the fact that the probability of fixation of beneficial alleles would have been higher on chromosomes with more frequent recombination even in the absence of epistasis.

SIMULATIONS

A Pascal program was written to test some of the above approximations and predictions. We focus on loci that are tightly linked since the effects are largest and contribute disproportionately to selection on a modifier. A diploid population of size 500,000 (or haploid population of size 1,000,000) was simulated using a three-locus model, where two loci (J and K) were subject to selection and the third locus (M) modified the recombination rate between loci J and K (recursions from FELDMAN 1972). For all runs, the initial frequency of the modifier allele that increased recombination, m_1, was $p_m = 0.5$. The beneficial allele, k_1, was set to a variety of initial frequencies at the time T when j_1 first appeared: 3.1×10^{-7} ($T = -15$), 4.5×10^{-5} ($T = -10$), 6.7×10^{-3} ($T = -5$), 7.6×10^{-2} ($T = -2.5$), 0.5 ($T = 0$), and 0.92 ($T = 2.5$). The modifier locus and locus K were assumed to be in linkage equilibrium when the beneficial mutation, j_1, occurred (further simulations indicated that the results are not sensitive to this assump-

tion). j_1 first appeared in a single copy at frequency 10^{-6} on a randomly chosen genetic background.

While j_1 was rare, a semi-stochastic simulation was performed. Every generation each chromosome carrying the new beneficial allele would have a number of offspring drawn from a Poisson distribution with a mean equal to the expected number of offspring based on the deterministic equations. The chromosomes not containing the new allele were allowed to change deterministically. The new allele was tracked until it either disappeared or rose to $7/s$ copies, in which case its future trajectory was assumed to be deterministic. [The probability of fixation of a beneficial allele that is at frequency $7/(2Ns)$ is $\approx 1 - 10^{-6}$, or very nearly one (CROW and KIMURA 1970, Equation 8.8.3.13).] When the two selected loci were fixed (less than one chromosome in the population was expected to carry the disappearing alleles), the frequency of the m_1 allele at the modifier locus was determined. This procedure was repeated 5,000,000 times for each parameter set tested and the average frequency of the modifier determined. The simulation results are presented in Figures 4–9. We assumed multiplicative selection across loci to generate the figures, although graphs with additive selection are quite similar (data not shown). Recall that under multiplicative selection, MAYNARD SMITH (1968) demonstrated that there would be no selection on recombination in a deterministic model with no initial disequilibrium.

Several points deserve mention. First, although the change in frequency at the modifier locus appears very small (*e.g.*, $\approx 10^{-4}$ for $T < 0$ in Figure 4), this change is averaged over both the rare occasions when j_1 fixed in the population and the vast majority of times when it was lost. Second, it should be noticed that the effect of the modifier was assumed to be small ($\delta r_2 = 0.01$) to provide an accurate comparison with the analysis. Proportionally larger changes in p_m are expected to occur with stronger modifiers of recombination. Third, the change in frequency at the modifier locus was highly variable, depending strongly on the genetic background on which j_1 appeared and on whether fixation of j_1 occurred. This variability makes it difficult to assess the success of the predictions. As a rough measure, note that the predicted results fell within ± 1.96 SE in 23 out of 36 cases tested.

The discrepancies observed have several potential sources. We have assumed that the modifier locus changes the probability of fixation of the new beneficial mutation but that it otherwise exerts no influence on the dynamics of the system. We know, however, that this assumption does not hold. The initial disequilibrium that is created between the selected loci creates indirect selection on the modifier locus as in the case where disequilibrium is generated by epistasis (BARTON 1995a; BERGMAN *et al.* 1995). When j_1 appears with k_1, there is initially positive disequilibrium and the modifier allele that decreases recombination increases in frequency. Similarly, when j_1 appears with k_0, there is initially negative disequilibrium and the modifier allele that increases recombination is favored. Even when the average disequilibrium across all genetic backgrounds is zero, the average change in frequency at the modifier locus is not. The effect of the initial disequilibrium on the change in p_m is shown by the dotted lines in Figures 4–9, which also incorporate changes in p_m resulting from the modifier's effect on fixation probability. The dotted lines were found by (i) calculating the probability of fixation for each of the four backgrounds using a numerical evaluation of (7) – (10), (ii) determining, for each genetic background, the change in p_m in a purely deterministic simulation with allele j_1 initially present on one genetic background at frequency $1/$

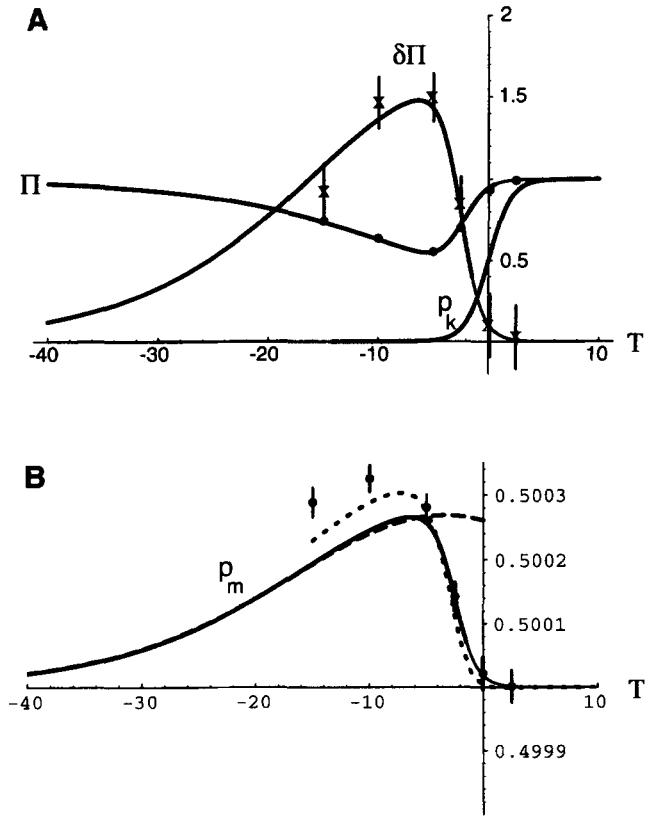

FIGURE 4.—Comparison of theoretical and simulation results for $r_1 = 0.001$, $r_2 = 0.02$, $\delta r_2 = 0.01$, $s = 0.01$, $S = 0.1$, and $2N = 10^6$. (A) Π, $\delta\Pi$ and p_k as a function of the time at which j_1 appears in the population. Dots give simulation results for Π, the average fixation probability scaled by the expectation of $2s$; error bars are too small to see. X's give simulation results for $\delta\Pi$, the effect of the modifier on the average fixation probability scaled by $2s\ \delta r_2/S$; error bars measure ± 1.96 SE. Solid lines give numerical solutions to (7) – (10). (B) Change in frequency at modifier locus. Dots give simulation results for the average frequency of the modifier upon fixation of the selected loci; error bars measure ± 1.96 SE. Solid line gives expectation based on (19) using a numerical evaluation of (7) – (10). Dashed line gives expectation based on (21). Dotted line estimates changes that occur at the modifier locus when modifier alleles are allowed to influence both the probability of fixation of beneficial mutants and the subsequent dynamics of linkage disequilibrium (see text).

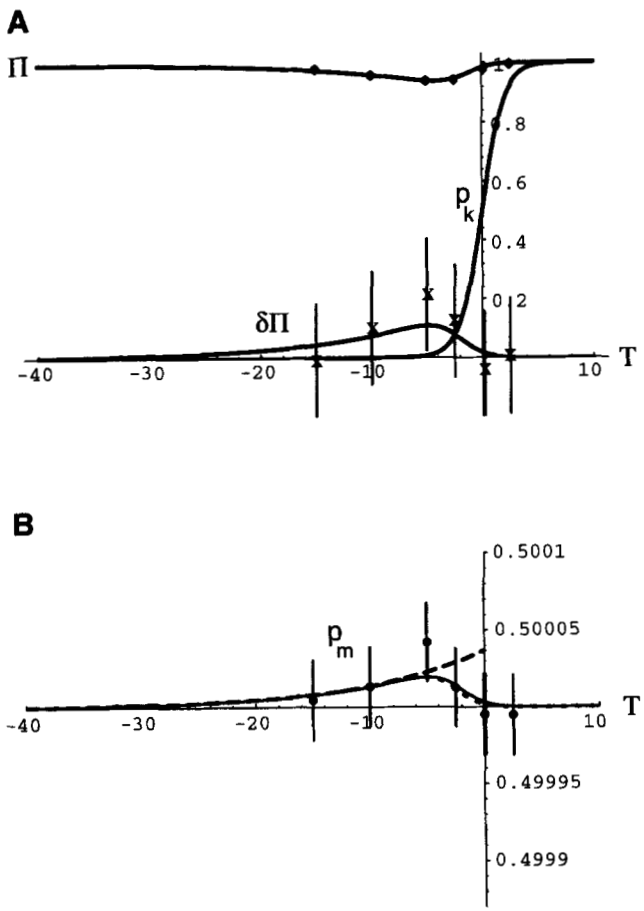

FIGURE 5.—Comparison of theoretical and simulation results for $r_2 = 0.1$, remaining terms are as in Figure 4.

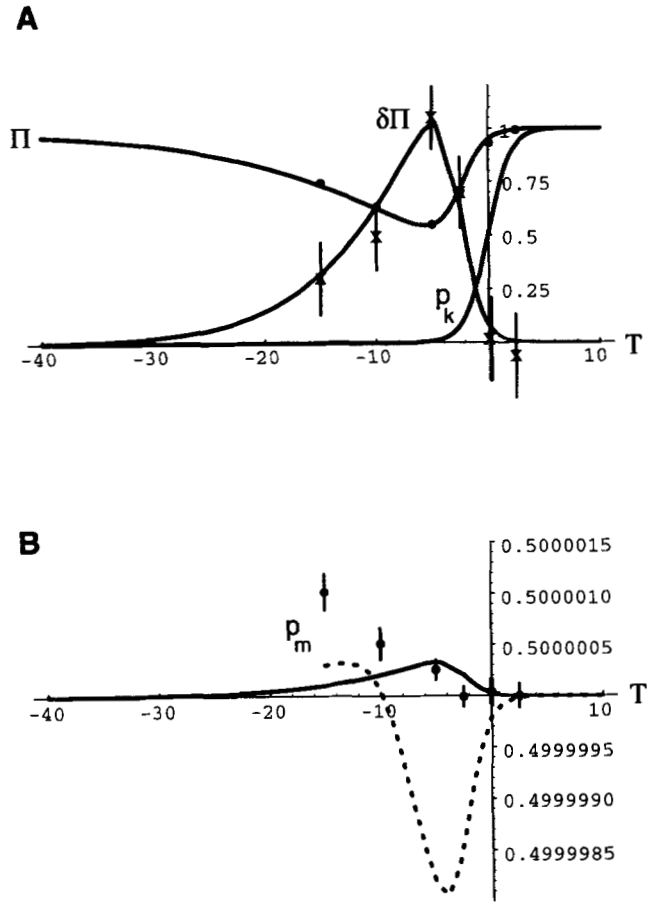

FIGURE 6.—Comparison of theoretical and simulation results for $r_1 = 0.01$, remaining terms are as in Figure 4.

$(4Ns)$ [we start the allele at $1/(2s)$ copies to account for the accelerated rise in allele frequency expected among processes that fix, discussed in APPENDIX C], and (iii) multiplying the expected frequency of each genetic background by (i) and (ii) and summing over all backgrounds. By this method, the change in frequency of the modifier can be estimated taking into account both the modifier allele's effect on the fixation probabilities and the response of the modifier to the initial disequilibrium between loci J and K. Randomly generated initial disequilibrium has an especially important influence on the dynamics of the modifier when selection is strong relative to the recombination rate, r_2, between the selected loci (as in Figures 7 and 9, where the dotted lines provide a better approximation to the simulations than the solid lines). Under these conditions, strong selection for the new mutant magnifies the initial linkage disequilibrium faster than recombination dissipates it, leading the modifier to evolve in response (N. H. BARTON and S. P. OTTO, unpublished results). This effect can be seen whether disequilibrium is generated due to the chance occurrence of a mutation on a particular haplotype or due to random genetic drift in a finite population. In either case, randomly generated disequilibria tend to favor increased recombination even in the absence of epistatic interactions among loci. This effect is explored by BARTON and OTTO (unpublished results) in a study of the dynamics of randomly generated linkage disequilibria.

Finally, it should be mentioned that there is an additional factor that may play a role in the simulation results. During the first few generations in which stochastic factors are important, those trajectories that just happen to make the disequilibrium between loci J and K more positive (coupling the beneficial alleles j_1 and k_1) are more likely to be those processes in which j_1 successfully fixes within the population. As a consequence, the disequilibrium between J and K among those processes that happen to fix will on average be more positive than the disequilibrium observed in those processes in which j_1 will be lost. This potential complication may be quite small, however, because, if r_2 is large relative to selection, the modifier will change little in response to the initial disequilibrium and, if r_2 is small relative to selection, recombination is unlikely to occur during the initial stochastic period and so differences in disequilibria between trajectories of the stochastic process may be minor.

DELETERIOUS MUTATIONS

Interference among selected loci also occurs when selection acts against deleterious mutations. As with di-

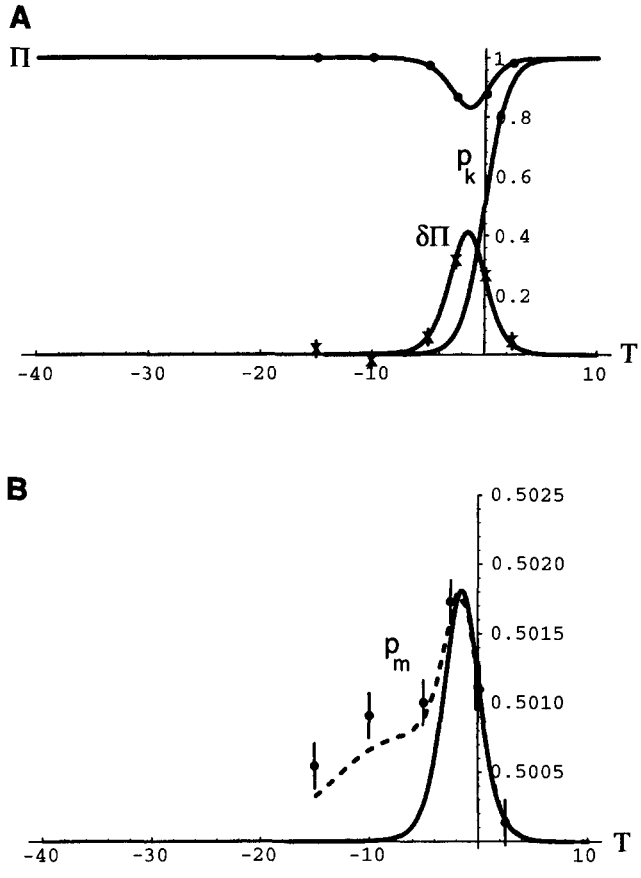

FIGURE 7.—Simulation results for $s = 0.1$, remaining terms are as in Figure 4. Dashed line is no longer present since $\theta = 1$ and (21) can no longer be used.

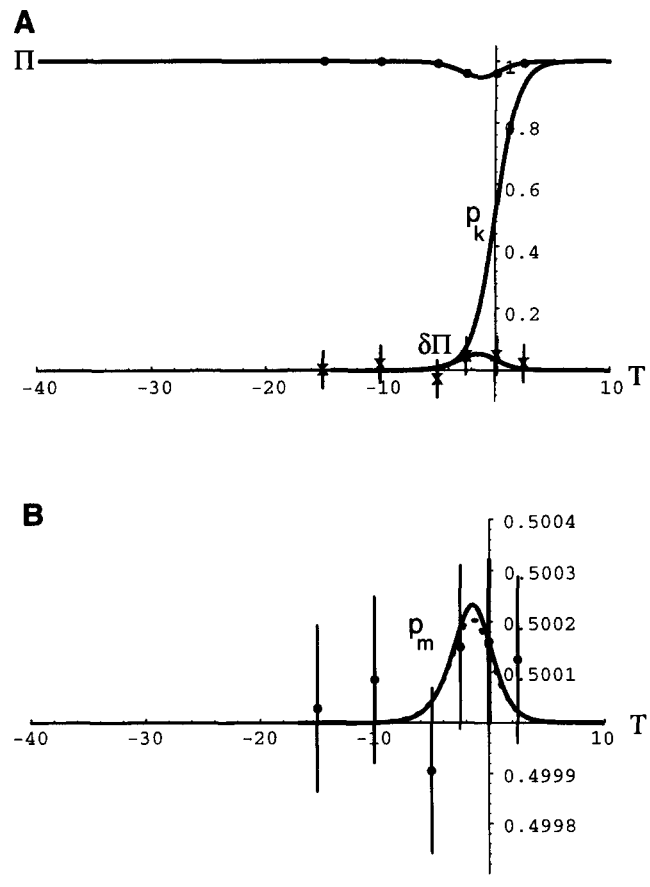

FIGURE 8.—Comparison of theoretical and simulation results for $r_2 = 0.1$ and $s = 0.1$, remaining terms are as in Figure 4. Dashed line is no longer present since $\theta = 1$.

rectional selection, purifying selection, on average, reduces the effectiveness of selection at linked loci. Therefore, new beneficial mutations have a lowered chance of fixation when there are neighboring loci held at a mutation-selection balance (BARTON 1995b). The effect of background selection against deleterious mutations is potentially important since many loci can contribute. In this section, we determine the extent to which a modifier of recombination can increase the probability of fixation of a beneficial mutation when selection is acting at other loci to eliminate deleterious mutations.

The model is a straightforward extension of the model used above. We assume that deleterious mutations occur at locus K at rate μ ($k_1 \rightarrow k_0$) and that selection acts against them with strength S as before. When a particular chromosome carrying the new beneficial allele j_1 reproduces, new deleterious mutations may occur at locus K thereby changing the genetic background of j_1. Hence, we must modify the branching process described by (1) and (2) to incorporate these new deleterious mutations. $P^*_{11}[t+1]$ (the expected fixation probability of a single chromosome bearing j_1 in generation $t+1$ that is the descendant of an $m_1 j_1 k_1$ chromosome in the previous generation), for example, becomes the following:

$$
\begin{aligned}
P^*_{11}[t+1] = {} & (1-\mu)(p_m p_k^A \\
& + p_m q_k^A(1-(r_2 + 2\,\delta r_2\, q_m)) + q_m p_k^A(1-r_1) \\
& + q_m q_k^A(1-r_1)(1-(r_2 - \delta r_2(p_m - q_m))))P_{11}[t+1] \\
& + \mu(p_m p_k^A + p_m q_k^A(1-(r_2 + 2\,\delta r_2\, q_m)) \\
& + q_m p_k^A(1-r_1) + q_m q_k^A(1-r_1) \\
& \times (1-(r_2 - \delta r_2(p_m - q_m))))P_{10}[t+1] \\
& + (p_m q_k^A(r_2 + 2\,\delta r_2\, q_m) + q_m q_k^A(1-r_1) \\
& \times (r_2 - \delta r_2(p_m - q_m)))P_{10}[t+1] \\
& + (1-\mu)(q_m p_k^A r_1 + q_m q_k^A r_1(1-(r_2 \\
& - \delta r_2(p_m - q_m))))P_{01}[t+1] \\
& + \mu(q_m p_k^A r_1 + q_m q_k^A r_1(1- \\
& (r_2 - \delta r_2(p_m - q_m))))P_{00}[t+1] \\
& + (q_m q_k^A r_1(r_2 - \delta r_2(p_m - q_m)))P_{00}[t+1], \quad (34)
\end{aligned}
$$

which takes into account all possible changes that may occur in the genetic background.

We again work in terms of the scaled variables (as defined in Table 1). In this case, however, we assume that the allele frequency of k_1 is held constant over time at a

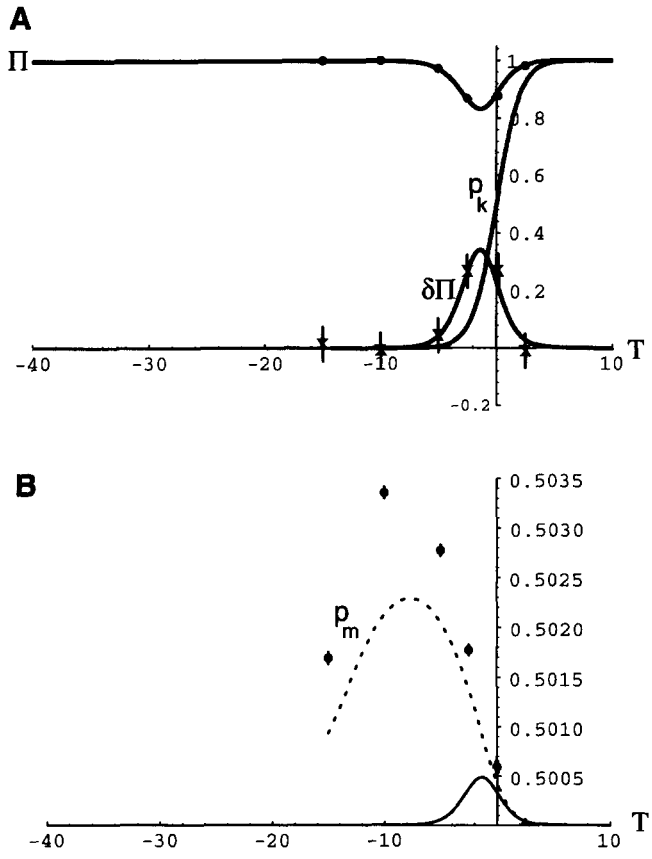

FIGURE 9.—Comparison of theoretical and simulation results for $r_1 = 0.01$ and $s = 0.1$, remaining terms are as in Figure 4. Dashed line is no longer present since $\theta = 1$.

mutation-selection balance and that therefore the probability of fixation of j_1 is independent of time. That is, $d\Pi/dT = 0$. After a bit of algebra, it can be found that

$$\frac{d\Pi}{dT} = -\theta\Pi(1 - \Pi) + \theta p_k q_k \Delta^2 = 0 \quad (35)$$

$$\frac{d\Delta}{dT} = \Delta(\rho_2 + (2\Pi - 1)\theta + (p_k - q_k)(1 - \theta\Delta)) - \Pi + \Delta q_k = 0. \quad (36)$$

$$\frac{d\delta\Pi}{dT} = (\rho_1 + \theta(2\Pi - 1))\delta\Pi + 2\theta p_k q_k \Delta\, \delta\Delta = 0 \quad (37)$$

$$\frac{d\delta\Delta}{dT} = (\rho_1 + \rho_2 - \chi + \theta(2\Pi - 1))\delta\Delta + (1 - 2\theta\Delta)((p_k - q_k)\delta\Delta - \delta\Pi) + \Delta + \delta\Delta\, q_k = 0. \quad (38)$$

These equations are approximate, having kept only leading order terms in the selection coefficients and recombination rates, and assume that recombination rates are not much greater than the selection coefficients.

Equations 35 and 36 were solved by BARTON (1995b), who analyzed the model without a modifier of recombination (note, though, that BARTON uses a different ordering of the alleles so that his Δ has the opposite sign). Assuming that the mutation rate is small so that the frequency of mutant alleles (q_k) is small, BARTON found that

$$\Pi = 1 - q_k\Delta^2 + O(q_k^2) \quad (39)$$

$$\Delta = \frac{(1 + \rho_2 + \theta) + \sqrt{(1 + \rho_2 + \theta)^2 - 4\theta}}{2\theta} + O(q_k). \quad (40)$$

From (37) and (38),

$$\delta\Pi = -\frac{2\theta q_k\Delta\, \delta\Delta}{\rho_1 + \theta} + O(q_k^2) \quad (41)$$

$$\delta\Delta = -\frac{\Delta}{\rho_1 + \rho_2 - \chi + \theta + 1 - 2\theta\Delta} + O(q_k). \quad (42)$$

Therefore the average effect of the modifier on the fixation probability of j_1 is given by

$$\delta\Pi = -\frac{2\theta q_k\Delta^2}{(\rho_1 + \theta)(\rho_1 + \rho_2 - \chi + \theta + 1 - 2\theta\Delta)} + O(q_k^2) = \frac{q_k S\left(r_2 + s + S - \sqrt{r_2^2 + 2r_2 s + s^2 + 2r_2 S - 2sS + S^2}\right)^2}{2s(r_1 + s)\left(r_1 + \sqrt{r_2^2 + 2r_2 s + s^2 + 2r_2 S - 2sS + S^2}\right)} + O(q_k^2). \quad (43)$$

For loose recombination, the derivation of (35)–(38) must be repeated to include terms such as $r_2 = \rho_2 S$ that are no longer negligible. [The statement on p. 830 of BARTON (1995b) that the equations for tight linkage extend to loose linkage was incorrect, since it did not allow for the difference between allele frequencies before and after selection.] After a lot of tedious bookkeeping, it can be shown, however, that (39)–(42) provide good estimates even under loose recombination. The effect of a modifier on the average probability of fixation is therefore given by (43) for all values of recombination. Equation 43 may be simplified further if recombination rates are either much lower or much higher than the selection coefficients. When recombination rates are negligibly low, we have

$$\Delta \approx 1 \quad \Pi \approx 1 - q_k \quad \delta\Pi \approx \frac{2q_k S}{S - s} \quad \text{for } s \ll S \quad (44)$$

$$\Delta \approx \frac{S}{s} \quad \Pi \approx 1 - \frac{q_k S^2}{s^2} \quad \delta\Pi \approx \frac{2q_k S^3}{s^2(s - S)} \quad \text{for } s \gg S. \quad (45)$$

[Equation 17b of BARTON (1995b) provided only the first solution for $s \ll S$ and should be supple-

mented by the above for $s \gg S$.] When recombination is loose,

$$\Delta \approx \frac{S}{s + S + r_2}$$

$$\Pi \approx 1 - q_k\left(\frac{S}{s + S + r_2}\right)^2 \quad \delta\Pi \approx \frac{2q_k s S^3}{r_1 r_2^2 (r_1 + r_2)}. \quad (46)$$

The probability of fixation of j_1 when it initially appears linked to m_1 (F_1) or to m_0 (F_0) is given by (11) and (12), respectively, using (43) for $\delta\Pi$. These equations demonstrate that (1) selection against deleterious mutations reduces the fixation probability of new beneficial mutations and (2) new beneficial mutations will have a higher probability of fixation if they arise with a modifier allele that increases recombination. Both of these effects are strongest when linkage is tight and decrease rapidly as recombination rates rise. Both effects are also proportional to the frequency of deleterious alleles at the selected locus, K, and so are small when considering the effects of a single mutation-selection balance on the fixation probability of a new beneficial mutation.

To find the expected change in frequency of the modifier locus, we again have to take into account the fact that increasing the probability of fixation makes it more likely that m_1 will hitchhike up in frequency. The amount of this hitch depends strongly on the recombination rate between the modifier and the new beneficial mutation. The average change in frequency of m_1, Δp_m, is given by (19) and (20) for tight and loose linkage, respectively, giving

$$\Delta p_m = \frac{2s\,\delta r_2}{S} p_m q_m (4Ns)^{-r_1/s} \Gamma\left(1 + \frac{r_1}{s}\right)^2 \Gamma\left(1 - \frac{r_1}{s}\right) \times \frac{q_k S\left(r_2 + s + S - \sqrt{r_2^2 + 2r_2 s + s^2 + 2r_2 S - 2sS + S^2}\right)^2}{2s(r_1 + s)\left(r_1 + \sqrt{r_2^2 + 2r_2 s + s^2 + 2r_2 S - 2sS + S^2}\right)}, \quad (47)$$

for tight linkage, and

$$\Delta p_m = \frac{2s\,\delta r_2}{S} p_m q_m \frac{1}{4N(r_1 - s)} \times \frac{q_k S\left(r_2 + s + S - \sqrt{r_2^2 + 2r_2 s + s^2 + 2r_2 S - 2sS + S^2}\right)^2}{2s(r_1 + s)\left(r_1 + \sqrt{r_2^2 + 2r_2 s + s^2 + 2r_2 S - 2sS + S^2}\right)} \quad (48)$$

for loose linkage. These reduce to an interpretable form if we assume that recombination is extremely tight (so that r_1 and r_2 terms can be ignored), in which case,

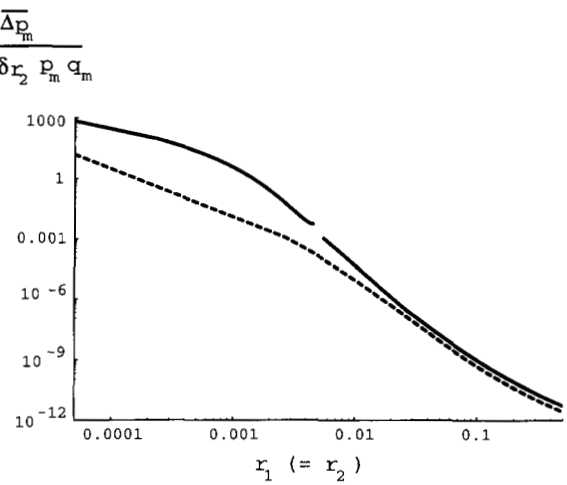

FIGURE 10.—The expected change at a modifier locus when multiple beneficial mutations occur throughout the population at a rate λ per genome per generation. Δp_m divided by $\Pi^*\lambda^2\delta r_2 p_m q_m$ is shown for increasing rates of recombination, using (31) (---) to approximate (27) for tight linkage and (29) for loose linkage (—). $s = 0.005$, $S = 0.01$, $r_1 = r_2$, and $2N = 10^6$. Note the axes are both on a log scale.

$$\Delta p_m = \frac{4s\,\delta r_2\, q_k p_m q_m}{S - s} \quad \text{for } s \ll S$$

$$\Delta p_m = \frac{4S^2\,\delta r_2\, q_k p_m q_m}{s(s - S)} \quad \text{for } s \gg S$$

and extremely tight linkage, (49)

or if we assume that linkage is very loose (r_1, $r_2 \gg s$, S),

$$\Delta p_m = \frac{s^2 S^2\,\delta r_2\, q_k p_m q_m}{N r_1 r_2^2 (r_1 + r_2 - r_1 r_2)(r_1 - s)}$$

for extremely loose linkage. (50)

At best, with complete linkage, the strength of selection on the modifier ($\Delta p_m/(p_m q_m)$) is extremely weak, $O(\delta r_2 q_k)$, since only deleterious mutations at a single locus are considered. We must hence consider the impact of multiple deleterious mutations.

Graphs of (47) and (48) show that (47) decreases rapidly with increasing recombination but becomes invalid when $r_1 \approx s$. When $r_1 > s$, (48) continues to decline with recombination at about where (47) left off. An approximation to (48),

$$\Delta p_m = \frac{2s\delta r_2}{S} p_m q_m \frac{1}{4Nr_1} \times \frac{q_k S\left(r_2 + s + S - \sqrt{r_2^2 + 2r_2 s + s^2 + 2r_2 S - 2sS + S^2}\right)^2}{2sr_1\left(r_1 + \sqrt{r_2^2 + 2r_2 s + s^2 + 2r_2 S - 2sS + S^2}\right)}, \quad (51)$$

provides a function that is valid for all r_1, r_2. This func-

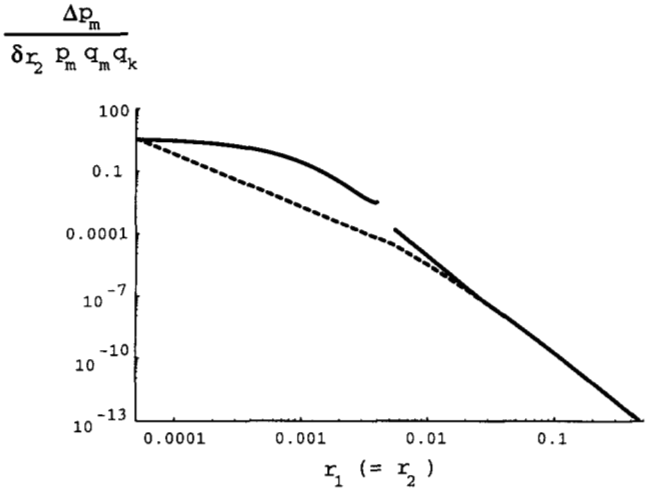

FIGURE 11.—The expected change at a modifier locus when both deleterious and beneficial mutations occur throughout the genome. Δp_m divided by $\delta r_2 p_m q_m q_k$ is shown for increasing rates of recombination, using (51) (---) to approximate (47) for tight linkage and (48) for loose linkage (—). $s = 0.005$, $S = 0.01$, $r_1 = r_2$, and $2N = 10^6$. Note the axes are both on a log scale.

tion provides a good estimate for Δp_m when recombination rates are high and an underestimate for Δp_m when recombination rates are low unless recombination is near zero (Figure 11). We can therefore use this function to determine the expected change at a modifier locus when loci are distributed along a chromosome. Let λ measure the rate of appearance of new beneficial mutations on the chromosome per generation and let n measure the number of loci at a mutation-selection balance on the chromosome. As before, we assume that there is a chromosome of length R morgans and that this chromosome is either circular or long enough that edge effects are relatively minor. The average change per generation in a modifier allele that increases recombination can then be found by evaluating the integrals in (30). Keeping only leading order terms in the selection coefficients and assuming that each of the n loci are equivalent, we have

$$\overline{\Delta p_m} \approx \lambda n \frac{\delta r_2\, p_m q_m q_k}{2N}\left[\left(\frac{1}{2R}\right)^2 \frac{\text{Max}\,[s, S]^2}{r_{small}} + \left(1 - \frac{1}{2R}\right)\frac{1}{2R}\left(\frac{4s^2S^2}{r_{small}} + \frac{4s^2S^2}{\text{Max}\,[s, S]}\right) + \left(1 - \frac{1}{2R}\right)^2 8s^2S^2\right]. \quad (52)$$

For reasonable values of the parameters, the largest contribution will always come from the first term in parentheses, that is, from linked loci. Directly comparing the first term of (32) for selection arising from interference among beneficial mutations to the first term of (52) for selection arising from interference between deleterious mutations and a beneficial mutation, the latter will only have a stronger effect if the total number of mutant alleles present on the chromosome is high: $n\, q_k > 2\Pi^*\lambda$.

To evaluate these results, we use the same rough estimates for chromosomes of higher eukaryotes drawn from data on humans and Drosophila (GRIFFITHS *et al.* 1996, pp. 136, 192, 494): $R \approx 1$, $r_{small} \approx 3 \times 10^{-4}$, and $S \approx s \approx 0.01$. To estimate n, the number of loci per chromosome at a mutation-selection balance note that, in humans, there are about 5000 genes per chromosome, for Drosophila, about 3000. We therefore take n to be 4000. With a per gene mutation rate of 10^{-5} and with $s \approx 0.01$, $q_k \approx 10^{-3}$. Using these data as a guide, an extremely rough estimate for the strength of selection favoring recombination at a modifier locus as a consequence of an increased fixation probability of beneficial mutations that appear on a chromosome with many other loci at a mutation-selection balance is

$$s_m \approx \frac{\overline{\Delta p_m}}{p_m q_m} \approx \frac{\lambda \delta r_2}{6N}. \quad (53)$$

If $\lambda \delta r_2 < 1$, selection on the modifier is extremely weak ($< 1/N$) and will be overwhelmed by drift. If $\lambda \gg 1$ such that $\lambda \delta r_2 > 1$, then selection for recombination due to interference between a locus with a new beneficial mutation and loci at a mutation-selection balance, (53), will tend to be weaker than selection for recombination due to interference between positively selected loci, (33), although the selective effects will act in concert. Of course, the conclusion that interference from deleterious alleles will affect the evolution of recombination less than interference from other beneficial alleles is only as good as the values used to estimate the parameters and would be incorrect if nq_k is much higher than our estimate of 4.

As with the case of multiple beneficial mutations (see Equation 27), selection for increased recombination is much stronger in organisms with infrequent recombination. In the extreme case of a completely congealed chromosome with n loci subject to deleterious mutations, we can use (49) to estimate the strength of selection on a modifier:

$$s_m \approx \frac{\overline{\Delta p_m}}{p_m q_m} \approx 4\lambda n\, \delta r_2\, q_k\, \text{Max}\left[\frac{s}{S - s}, \frac{S^2}{s(s - S)}\right]. \quad (54)$$

With the parameters used above, the strength of selection for increased recombination would be roughly on the order of $\lambda\, \delta r_2$ per generation, which indicates that as long as adaptive mutations occur there will be selection for recombination rates to increase. It is possible, when λ is low, for (54) to be larger than (27) indicating that, in a congealed genome, recombination may evolve to reduce the impact of loci at a mutation-selection balance on the probability of fixation of new beneficial mutations rather than to reduce the interference among multiple beneficial alleles.

DISCUSSION

One of the oldest hypotheses for the advantage of recombination is that it allows beneficial alleles that initially appear on separate chromosomes to be placed together onto the same chromosome (MORGAN 1913; FISHER 1930; MULLER 1932). In the absence of recombination, the success of a new mutation is limited by the fate of the chromosome on which it arises. In this case, the dynamics of the mutant allele are only weakly governed by its own selective advantage. The mutant allele will tend to drift rapidly up or down in frequency depending on the chromosome on which it appears. In essence, linkage with other selected loci reduces the effectiveness of selection over drift on the dynamics of the new beneficial allele. An important consequence is that the fixation probability of the new beneficial allele is reduced from the expectation of $2s$. Thus linked loci interfere with each others progress and place a limit on the response of a population to selection, a phenomenon known as the Hill-Robertson effect (HILL and ROBERTSON 1966). With recombination, the fates of alleles at different loci become uncoupled, and selection can more directly and effectively lead to the increase in frequency of a beneficial allele. Recombination thus reduces the limits placed on selection by interference among selected loci and allows populations to evolve at a higher rate.

Despite the age and appeal of this hypothesis, modifier models have never been developed to follow the evolution of recombination in the presence of new beneficial mutations that may or may not fix, although the effect has been studied by simulation (FELSENSTEIN and YOKOYAMA 1976; CHARLESWORTH *et al.* 1977). To determine whether the Hill-Robertson effect can select for increased recombination at modifier loci requires the analysis of a dynamical system that is not at equilibrium and that is subject to stochastic effects. Using a branching process model to describe the changes that occur over time in the fixation probability of an allele, BARTON (1995b) was able to show that recombination does increase the fixation probability of new beneficial alleles when there is selection at linked loci. Here we use similar methods to show that increased recombination can be favored at a modifier locus, even though the loci under selection do not interact epistatically. The main assumptions of the model are that selection acts multiplicatively upon two loci (J and K), that the population size is large, that selection is relatively weak so that a continuous time approximation can be made, and that the modifier alleles cause only slight differences in recombination rates. Our analysis was based on the estimation of two quantities: the average fixation probability (F_i) of a new beneficial mutation j_1 linked to a modifier allele m_i and the subsequent change in frequency of the modifier allele m_i due to hitchhiking with the beneficial mutation.

F_i was measured by averaging the fixation probability of j_1 when initially linked with either of the two alleles (k_0 and k_1) at the second selected locus (the average being weighted by the frequency of the two alleles). Letting m_1 represent a modifier allele that increases recombination by an amount δr_2 between the selected loci, the fixation probability of j_1 when it is initially linked to m_1 is greater than when it is initially linked with m_0 by an amount: $F_1 - F_0 = 2(s/S)\delta r_2\, \delta\Pi$ (see Table 1 for definitions). $\delta\Pi$ and hence F_1 and F_0 can be found numerically for any case by solving for the fixation probabilities given by (1) and its analogues. Analytical approximations were obtained for $\delta\Pi$ either at a given time point assuming tight linkage (13), or integrated over all possible times with tight linkage (58) or with loose linkage (75). As illustrated in Figure 2, $\delta\Pi$ (and hence the effect of the modifier, $F_1 - F_0$) is maximal (1) when the beneficial mutation j_1 arises while the beneficial allele at the second selected locus, k_1, is still uncommon ($T < 0$), (2) when recombination between the selected loci is rare, and (3) when recombination between the modifier and the selected loci is also rare. This last condition is less obvious and deserves some explanation. Tighter linkage between the modifier and the selected loci increases the difference in fixation probability between modifier backgrounds ($F_1 - F_0$) because the beneficial allele stays coupled with the modifier allele with which it arose for a longer period of time, thereby maintaining recombination at the rate determined by its initial genotype at the modifier locus.

The second step in our derivation involved estimating the expected change in the modifier allele m_1 through the process of hitchhiking with a beneficial allele. In APPENDIX C, we obtained estimates for the expected change in frequency of m_1 due to hitchhiking when a beneficial allele destined to fix first arises with m_1 ($\Delta_1 p_m$) or with m_0 ($\Delta_0 p_m$). Importantly, the hitch ($\Delta_i p_m$) is roughly proportional to $1/r_1$ and decreases rapidly with increasing recombination between the modifier and the selected locus; it also decreases with increasing population size. Altogether, the change in frequency of the modifier allele is given by $\overline{\Delta p_m} = p_m \Delta_1 p_m F_1 + q_m \Delta_0 p_m F_0$, where $\overline{\Delta p_m}$ estimates the total expected change at the modifier locus per new beneficial allele that arises. Since $\Delta_0 p_m = -(p_m/q_m)\Delta_1 p_m$ (APPENDIX C), this simplifies further to $\overline{\Delta p_m} = p_m \Delta_1 p_m (F_1 - F_0)$, where $\Delta_1 p_m$ is positive. Therefore, a modifier allele that increases recombination, say m_1, rises in frequency by this process since it increases the fixation probability of new mutations ($F_1 - F_0 > 0$). Using the estimates based on the three-locus model, we generalized our results by taking into account beneficial mutations that occur throughout a chromosome at a particular rate (32).

Finally, we considered the evolution of recombina-

tion when beneficial mutations arise on a chromosome bearing multiple deleterious mutations. We find that selection against deleterious mutations also interferes with the selective progress of a new beneficial allele and consequently reduces its fixation probability. As was the case with beneficial mutations, when there are loci at a mutation-selection balance, modifiers that increase recombination increase the fixation probability of new beneficial mutations and thereby gain a hitch that increases their frequency. Even though increased recombination rates are favored with either directional selection or purifying selection at other loci, we found that purifying selection against deleterious mutations tends to exert less of a selective pressure on modifiers of recombination, although this conclusion is based on extremely rough estimates of mutation rates and selection coefficients.

The main result of this article is that there is always selection for increased recombination at modifier loci as a result of the fact that recombination increases the fixation probability of favorable alleles; this selection can be strong, but only when loci are tightly linked or when the rate of appearance of new beneficial mutations is high. In the absence of recombination, selection at modifier loci strongly favors increased recombination both because a little recombination leads to a large increase in the fixation probability of beneficial mutations and because the extent of hitchhiking is maximal. Once relatively low levels of recombination have evolved (perhaps as low as one crossover event per chromosome), the selective pressure for further increases in the rate of recombination is drastically reduced and will be negligible compared to drift if the rate of new beneficial mutations is low (as might occur in a stable environment). There are three reasons why the strength of selection for recombination decreases rapidly with increasing recombination. (1) With some recombination, there is less potential benefit from further increases in recombination since the fixation probability of a beneficial allele is less influenced by selection at other loci. (2) With recombination between the modifier locus and the new beneficial mutation, the fixation probability of the mutation depends less on which modifier allele it arises with and more on the average recombination rate. (3) Finally, the extent of hitchhiking is reduced by recombination between the modifier and the selected loci. Because these three effects act in concert, selection for increased recombination at the modifier locus essentially disappears once there is moderate recombination if the environment is relatively stable, even though further increases in recombination would continue to raise the fixation probability of beneficial mutations. In a rapidly changing environment, however, adaptive mutations may appear at a high enough rate with strong enough selection acting upon them that there would again be substantial selection for increased recombination.

This leads us to conclude that the evolution of recombination may be a much more dynamic process than previously thought. During periods of rapid evolutionary change, higher rates of recombination would be favored since modifiers that increase recombination increase the fixation rate of adaptive mutations. Indeed, substantial changes in recombination rates have been observed among closely related species (reviewed by BROOKS 1988). For example, TRUE *et al.* (1996) found that recombination rates were 1.8 times higher in *D. mauritiana* (an island endemic) than in *D. melanogaster* and suggested that this increased recombination may have evolved in response to intense selection in a novel environment. During relatively static periods, however, selection for increased recombination would become negligible as long as some low level of recombination were occurring. During such periods, other selective forces, such as the physiological constraints discussed in the introduction, might govern the observed level of recombination.

In our view, this synthetic hypothesis for the evolution of recombination is most appealing: recombination has both immediate selective benefits to the individual as well as longer term benefits due to the production of variable offspring. The relative importance of these different selective forces on recombination will depend on the speed of change in the physical and biotic environment of an organism, with bursts of recombination occurring during bursts of evolutionary change. This hypothesis acknowledges the constraints imposed on recombination rates by the need for proper disjunction at meiosis (BAKER *et al.* 1976; HAWLEY 1988), but also explains why recombination rates would increase after periods of strong directional selection both in selection experiments (reviewed by KOROL and ILIADI 1994) and under artificial selection (BURT and BELL 1987). As an example, we found that the changes in recombination rates observed in a selection experiment for geotaxis in Drosophila (KOROL and ILIADI 1994) are at least consistent with the selection coefficients for recombination estimated from our model. Epistatic interactions could also potentially explain the evolution of increased recombination in such cases (BARTON 1995a), but only by recourse to the additional assumption that alleles affecting the trait under selection have negative epistatic interactions. Further experiments are needed to settle the issue of whether the Hill-Robertson effect or epistatic interactions drive the evolution of recombination that has been observed during periods of rapid evolutionary change.

Although we have focused entirely on the evolution of recombination, our results are also pertinent to an understanding of the evolution of sex. When sexual reproduction is rare, the fate of a beneficial mutation is once again limited by the genetic background on which it arises. Sex, with recombination, reduces the interference among selected loci, allowing selection to

act more directly upon the alleles at a locus. The methods developed within this paper can be used to examine the evolution of a modifier that affects the rate of sexual reproduction in a population. For a haploid population, the equations developed within this paper may be applied directly by setting r_i to the rate of recombination between two loci times the probability of sexual reproduction. A modifier that alters the probability that an organism reproduces sexually can then be studied by treating it as a modifier of recombination. For a diploid population, however, sex has an additional effect: sex allows chromosomes to reassort and hence unties the fate of one chromosome from the fate of its homologue even in the absence of recombination (KIRKPATRICK and JENKINS 1989). Reassortment of chromosomes and recombination within them, processes made possible by sexual reproduction, will decrease the interference between selected alleles, increase the fixation probability of new beneficial mutations, and consequently reduce the limits placed upon natural selection by linkage. The modifier analysis performed in this paper indicates that increased rates of recombination and sexual reproduction will be selectively favored at modifier loci, but that this selection will only be substantial when the level of recombination is low or when the environment is rapidly changing.

The authors thank ALAN BOULTON, BRIAN CHARLESWORTH, ANDY CLARK, ALEX KONDRASHOV, ROSIE REDFIELD, DOLPH SCHLUTER, and MICHAEL WHITLOCK for helpful discussions and/or comments on the manuscript. Their insights and assistance have added greatly to this work. This work was supported by a Science and Engineering Research Council grant (to N.B.), by the Darwin Trust of Edinburgh, and by a Natural Sciences and Engineering Research Council of Canada grant (to S.O.).

LITERATURE CITED

ABROMOWITZ, M., and I. A. STEGUN, 1970 *Handbook of Mathematical Functions.* Dover Publications, Inc., New York.

ALTENBERG, L., and M. W. FELDMAN, 1987 Selection, generalized transmission and the evolution of modifier genes. I. The reduction principle. Genetics **117:** 559–572.

BAKER, B. S., A. T. C. CARPENTER, M. S. ESPOSITO, R. E. ESPOSITO and L. SANDLER, 1976 The genetic control of meiosis. Annu. Rev. Genet. **10:** 53–134.

BARTON, N. H., 1994 The reduction in fixation probability caused by substitutions at linked loci. Genet. Res. **64:** 199–208.

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

BARTON, N. H., 1995b Linkage and the limits to natural selection. Genetics **140:** 821–841.

BASILE, G., M. AKER and R. K. MORTIMER, 1992 Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene *rad51.* Mol. Cell. Biol. **12:** 3235–3246.

BERGMAN, A., S. P. OTTO and M. FELDMAN, 1995 On the evolution of recombination in haploids and diploids I. Deterministic models. Complexity **1:** 49–57.

BERNSTEIN, H., F. HOPF and R. MICHOD, 1988 Is meiotic recombination an adaptation for repairing DNA, producing genetic variation, or both?, pp. 194–211 in *The Evolution of Sex, An Examination of Current Ideas,* edited by R. MICHOD and B. LEVIN. Sinauer Associates, Sunderland, MA.

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

BROOKS, L., 1988 The evolution of recombination rates, pp. 87–105 in *The Evolution of Sex: An Examination of Current Ideas,* edited by R. E. MICHOD and B. R. LEVIN. Sinauer Associates, Sunderland, MA.

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

BURT, A., G. BELL and P. H. HARVEY, 1991 Sex differences in recombination. J. Evol. Biol. **4:** 259–277.

CHARLESWORTH, B., 1976 Recombination modification in a fluctuating environment. Genetics **83:** 181–195.

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

CHARLESWORTH, B., and D. CHARLESWORTH, 1985 Genetic variation in recombination in *Drosophila.* I. Responses to selection and preliminary genetic analysis. Heredity **54:** 71–83.

CHARLESWORTH, B., I. MORI and D. CHARLESWORTH, 1985 Genetic variation in recombination in *Drosophila.* III. Regional effects on crossing over and effects on non-disjunction. Heredity **55:** 209–221.

CHARLESWORTH, D., B. CHARLESWORTH and C. STROBECK, 1977 Effects of selfing on selection for recombination. Genetics **86:** 213–226.

COX, M. M., 1993 Relating biochemistry to biology: how the recombinational repair function of RecA protein is manifested in its molecular properties. BioEssays **15:** 617–623.

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

CROW, J., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory.* Harper and Row, New York.

ENGEBRECHT, J., J. HIRSCH and G. S. ROEDER, 1990 Meiotic gene conversion and crossing over: their relationship to each other and to chromosome synapsis and segregation. Cell **62:** 927–937.

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

FELDMAN, M. W., F. B. CHRISTIANSEN and L. D. BROOKS, 1980 Evolution of recombination in a constant environment. Proc. Natl. Acad. Sci. USA **77:** 4838–4841.

FELDMAN, M. W., S. P. OTTO and F. B. CHRISTIANSEN, 1997 Population genetic perspectives on the evolution of recombination. Annu. Rev. Genet. **30:** 261–295.

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 University Press, Oxford.

GRIFFITHS, A. J. F., J. H. MILLER, D. T. SUZUKI, R. C. LEWONTIN and W. M. GELBART, 1996 *An Introduction to Genetic Analysis.* W. H. Freeman and Company, New York.

HALDANE, J. B. S., 1927 A mathematical theory of natural and artificial selection, Part V: Selection and mutation. Proc. Camb. Philos. Soc. **23:** 838–844.

HAWLEY, R. S., 1988 Exchange and chromosomal segregation in eucaryotes, pp. 497–528 in *Genetic Recombination,* edited by R. KUCHERLAPATI and G. R. SMITH. American Society for Microbiology, Washington DC.

HAWLEY, R. S., and W. E. THEURKAUF, 1993 Requiem for distributive segregation: achiasmate segregation in *Drosophila* females. Trends Genet. **9:** 310–317.

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

KARLIN, S., 1973 Sex and infinity: a mathematical analysis of the advantages and disadvantages of genetic recombination, pp. 155–194 in *The Mathematical Theory of the Dynamics of Biological Populations,* edited by M. S. BARTLETT and R. W. HIORNS. Academic Press, New York.

KIRKPATRICK, M., and C. D. JENKINS, 1989 Genetic segregation and the maintenance of sexual reproduction. Nature **339:** 300–301.

KNAPP, E., 1960 Tetrad analyses in green plants. Can. J. Genet. Cytol. **2:** 89–95.

KOEHLER, K. E., R. S. HAWLEY, S. SHERMAN and T. HASSOLD, 1996 Recombination and nondisjunction in humans and flies. Hum. Mol. Genet. **5:** 1495–1504.

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

KONDRASHOV, A. S., 1993 Classification of hypotheses on the advantage of amphimixis. J. Hered. **84:** 372–387.

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

KUSANO, K., Y. SUNOHARA, N. TAKAHASHI, H. YOSHIKURA and I. KOBAYASHI, 1994 DNA double-strand break repair: genetic determinants of flanking crossing-over. Proc. Natl. Acad. Sci. USA **91:** 1173–1177.

LI, W-H., and D. GRAUR, 1991 *Fundamentals of Molecular Evolution*. Sinauer Associates, Sunderland, MA.

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

MAYNARD SMITH, J., 1978 *The Evolution of Sex*. Cambridge University Press, Cambridge.

MAYNARD-SMITH, J., and J. HAIGH, 1974 The hitch-hiking effect of a favorable gene. Genet. Res. **23:** 23–35.

MERRIAM, J. R., and J. N. FROST, 1964 Exchange and nondisjunction of the *X* chromosomes in female *Drosophila melanogaster*. Genetics **49:** 109–122.

MICHOD, R., and B. LEVIN, 1988 *The Evolution of Sex, An Examination of Current Ideas*. Sinauer Associates, Sunderland, MA.

MORGAN, T. H., 1913 *Heredity and Sex*. Columbia University Press, New York.

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

MUKAI, T., S. I. CHIGUSA, L. E. METTLER and J. F. CROW, 1972 Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. Genetics **72:** 335–355.

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

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

ORR, H. A., and J. A. COYNE, 1992 The genetics of adaptation: a reassessment. Am. Nat. **140:** 725–742.

OTTO, S., and M. W. FELDMAN, 1997 Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theoret. Popul. Biol. **51:** 134–147.

STEPHAN, W., T. H. E. WIEHE and M. W. LENZ, 1992 The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. Theoret. Popul. Biol. **41:** 237–254.

TRUE, J. R., J. M. MERCER and C. C. LAURIE, 1996 Differences in crossover frequency and distribution among three sibling species of *Drosophila*. Genetics **142:** 507–523.

VILLENEUVE, A. M., 1994 A *cis*-acting locus that promotes crossing over between *X* chromosomes in *Caenorhabditis elegans*. Genetics **136:** 887–902.

WOLFRAM RESEARCH, I., 1993 *Mathematica, Version 2.2*. Wolfram Research, Champaign, IL.

Communicating editor: A. G. CLARK

APPENDIX A
APPROXIMATION UNDER TIGHT LINKAGE AND WEAK SELECTION AT LOCUS *J*

When selection is weak, the average fixation probability decreases slowly over time and returns to $2s$ while the substituting allele is rare (see Figures 1, 4–9). Similarly, the extent to which the modifier alters the probability of fixation increases from zero and returns again while p_k remains less than a half ($T < 0$, see Figure 1). Therefore, for weak selection on the new mutant, we can approximate $\delta\Pi$ by focusing on the period in which $T \ll 0$ and $p_k \ll 1$. During this period, $d\delta\Pi/dT$ is dominated by the first half of expression (9),

$$\frac{d\delta\Pi}{dT} \approx (\rho_1 + \theta(2\Pi - 1))\,\delta\Pi. \qquad (55)$$

Equation 55 may be integrated explicitly after substitution from (14) for Π, giving

$$\delta\Pi \approx \frac{e^{\rho_1 T} e^{\theta T} c}{(1 - \theta^{\rho_2} + \theta^{\rho_2} e^{\theta T})^2}, \qquad (56)$$

where c is a constant of integration. The following heuristic method provides an estimate for c by determining the expected value of $\delta\Pi$ when ρ_1 equals zero. When ρ_1 equals zero, the new beneficial mutation, j_1 will not recombine away from the modifier allele with which it arises. Therefore the probability of fixation of j_1 should depend on its genetic background at the modifier locus only to the extent that the fixation probability is sensitive to a change in the recombination rate in the absence of a modifier. That is, we expect that $\delta P_k = (\partial P_k/\partial r_2)\,\delta r_2$. Using the definitions of Π and $\delta\Pi$, this implies that $\delta\Pi|_{\rho_1=0} = \partial\Pi/\partial\rho_2$. Using (14) for Π (BARTON 1995b),

$$\frac{\partial\Pi}{\partial\rho_2} = \frac{e^{\theta T}\theta^{\rho_2}\log\left(\frac{1}{\theta}\right)}{(1 - \theta^{\rho_2} + \theta^{\rho_2} e^{\theta T})^2},$$

implying that $c = \theta^{\rho_2}\log(1/\theta)$ and

$$\delta\Pi \approx \frac{e^{\rho_1 T} e^{\theta T}\theta^{\rho_2}\log\left(\frac{1}{\theta}\right)}{(1 - \theta^{\rho_2} + \theta^{\rho_2} e^{\theta T})^2} = e^{\rho_1 T}\frac{\partial\Pi}{\partial\rho_2}. \qquad (57)$$

When ρ_1 is not equal to zero, the sensitivity of the fixation probability to recombination ($\partial\Pi/\partial\rho_2$) is discounted by the amount $e^{\rho_1 T}$. Presumably, $e^{\rho_1 T}$ measures how much recombination occurs between the modifier locus and the selected locus, J, before the fate of the new mutant j_1 is decided. When j_1 arises while k_1 is still rare, its fate is not determined until k_1 is nearly fixed and the amount of recombination between M and J in this period will be $e^{\rho_1 T}$. Recombination during this period imparts some of the advantage of increasing recombination between the selected loci onto the alternative allele at the modifier locus and reduces the effect of the modifier.

Equation 56 was compared to numerical solutions of the full system of differentials (7)–(10). Parameters chosen were each combination of $\rho_1 = \{0.001, 0.01, 0.1, 1, 10\}$, $\rho_2 = \{0.001, 0.01, 0.1, 1, 10\}$, and $\theta = \{0.001, 0.01, 0.1\}$. The approximation (57) was very good for $T < 0$ whenever $\rho_1 < 1$ and $\rho_2 < 1$. The approximation was also good for $\rho_1 < 1$ and $\rho_2 = 1$, but poor for the remaining parameter sets tested. This comparison is also shown in Figures 1 and 2 (compare dashed to solid lines for $\delta\Pi$).

The net effect of hitchhiking on a modifier of recombination can be found by integrating $\delta\Pi$ over all possible times of origin of new beneficial mutations. By substituting $z = e^T$ into (57) and by noting that $\delta\Pi$ is near zero for $T > 0$, we obtain

$$\int_{-\infty}^{\infty} \delta\Pi dT \approx \int_{-\infty}^{0} \delta\Pi dT = \int_0^1 \frac{z^{\rho_1} z^{\theta} \theta^{\rho_2} \log\left(\frac{1}{\theta}\right)}{z(1-\theta^{\rho_2}+\theta^{\rho_2} z^{\theta})^2}\, dz$$

$$= \frac{\log[\theta]\left(\rho_1 F\left[1, 1+\frac{\rho_1}{\theta}, 2+\frac{\rho_1}{\theta}, -\frac{\theta^{\rho_2}}{1-\theta^{\rho_2}}\right] - (\rho_1+\theta)(1-\theta^{\rho_2})\right)}{(\rho_1+\theta)\theta^{(1-\rho_2)}(1-\theta^{\rho_2})^2}, \quad (58)$$

where F is the hypergeometric function. Under special conditions, the above can be simplified further:

$$\int_{-\infty}^{0} \delta\Pi\, dT = \begin{cases} \dfrac{(\theta^{\rho_2} + \log(1-\theta^{\rho_2}))\log(\theta)}{\theta^{(1+\rho_2)}} & \text{for } \rho_1 = \theta \\ \dfrac{\log\left(\frac{1}{\theta}\right)(\theta^{(1+\rho_2)} + \rho_1 \log(1-\theta^{\rho_2}))}{\theta^2(1-\theta^{\rho_2})} & \text{for } \rho_1 \ll 1 \\ \dfrac{1}{\rho_2 \theta} & \text{for } \rho_1, \rho_2 \ll \theta. \end{cases} \quad (59)$$

Equation 58 is compared to numerical solutions of the differential equations in Table 3.

APPENDIX B
APPROXIMATION UNDER LOOSE LINKAGE

In (7)–(10), we assumed that terms such as $\rho_2 S$ were negligible. This is not a tenable assumption when recombination rates are high. With loose linkage, $\delta\Pi$ and $\delta\Delta$ are both very small and a good approximation for the scaled variables can be obtained by setting $\delta\Pi = O(\epsilon^2)$, $\delta\Delta = O(\epsilon^2)$, $\Pi = 1 - O(\epsilon)$, $\Delta = O(\epsilon)$, $s = O(\epsilon)$, and $S = O(\epsilon)$ and ignoring terms of $O(\epsilon^3)$. We then have

$$\frac{d\Pi}{dT} = -\theta\Pi(1-\Pi) + \theta p_k q_k \Delta^2 - r_2 p_k q_k \Delta(\Delta\theta(2-r_2) - \theta S - S p_k) + O(\epsilon^3) \quad (60)$$

$$\frac{d\Delta}{dT} = \Delta(\rho_2 + (2\Pi - 1)\theta + (p_k - q_k)(1-\theta\Delta)) - \Pi + r_2\Delta(\theta - (p_k - q_k)) + r_2\theta\Delta^2(p_k - q_k)(2-r_2) - 2r_2\theta\Pi\Delta + O(\epsilon^3). \quad (61)$$

$$\frac{d\delta\Pi}{dT} = (\rho_1 + \theta(2\Pi - 1))\delta\Pi + 2\theta p_k q_k \Delta\, \delta\Delta + r_1 p_k q_k \delta\Delta - r_1\theta\delta\Pi + O(\epsilon^3) \quad (62)$$

$$\frac{d\delta\Delta}{dT} = (\rho_1 + \rho_2 - \chi + \theta(2\Pi - 1))\delta\Delta + (1 - 2\theta\Delta)((p_k - q_k)\delta\Delta - \delta\Pi) + \Delta - (\theta + (p_k - q_k))(\delta\Delta(r_1 + r_2 - r_1 r_2) + S\Delta) + r_1\delta\Pi + O(\epsilon^3). \quad (63)$$

Equations 7 and 9 provide adequate solutions for Π and $\delta\Pi$ when θ is large, but can be off by orders of magnitude when θ is small in which case (60) and (62) must be used. In either case, (8) and (10) do provide adequate estimates for Δ and $\delta\Delta$ and can be used to determine the net effect of the modifier on fixation probability. The methodology that we will follow is similar to the one employed in Appendix B of BARTON (1995b).

When either recombination is loose or selection is strong, Π deviates only slightly from one and $\delta\Pi$ only slightly from zero. Let $\Pi = 1 - \epsilon_1$ and $\delta\Pi = \epsilon_2$. To leading order in ϵ_i we can write (62) as

$$\frac{d\epsilon_2}{dT} = (\rho_1 + \theta(1 - r_1))\epsilon_2 + 2\theta p_k q_k \Delta\delta\Delta + r_1 p_k q_k \delta\Delta \quad (64)$$

$$\epsilon_2 = -\int_T^{\infty} e^{-(\rho_1 + \theta(1-r_1))(\tau - T)}(2\theta p_k(\tau) q_k(\tau)\Delta\, \delta\Delta + r_1 p_k(\tau) q_k(\tau)\delta\Delta)\, d\tau. \quad (65)$$

The net effect of a substitution on $\delta\Pi$ can be found by integrating the above over all time.

$$\int_{-\infty}^{\infty} \delta\Pi\, dT = -\int_{-\infty}^{\infty}\int_T^{\infty} e^{-(\rho_1+\theta(1-r_1))(\tau-T)} 2\theta p_k(\tau) q_k(\tau)\Delta\, \delta\Delta\, d\tau dT - \int_{-\infty}^{\infty}\int_T^{\infty} e^{-(\rho_1+\theta(1-r_1))(\tau-T)} r_1 p_k(\tau) q_k(\tau)\delta\Delta\, d\tau dT. \quad (66)$$

We shall solve each of the two integrals of (66) in turn.

First part of (66): By reversing the order of integration, we get

$$-\int_{-\infty}^{\infty} 2\theta p_k(\tau) q_k(\tau)\Delta\, \delta\Delta \frac{1}{\rho_1 + \theta(1-r_1)}\, d\tau. \quad (67)$$

An estimate for Δ is given by Equation B4 of BARTON (1995b):

$$\Delta \approx \frac{1}{p_k(T) q_k(T)} \int_T^{\infty} p_k(\tau) q_k(\tau) e^{(\rho_2+\theta)(T-\tau)}\, d\tau.$$

To evaluate, $\delta\Delta$ we note that $\delta\Pi$ is small and may be ignored relative to $\delta\Delta$. Following BARTON (1995b), we will assume that $\theta\Delta$ is fairly small when selection is strong or recombination loose, indicating that $(1 - 2\theta\Delta)$ is approximately 1. We then have that

$$\frac{d\delta\Delta}{dT} \approx [\rho_1 + \rho_2 - \chi + \theta + p_k(T) - q_k(T)]\,\delta\Delta + \Delta.$$

$$\delta\Delta \approx -\frac{1}{p_k(T)\,q_k(T)} \times \int_T^\infty p_k(\tau)\,q_k(\tau)\,\Delta e^{(\rho_1+\rho_2-\chi+\theta)(T-\tau)}\,d\tau.$$

Substituting the above into (67), we obtain

$$\frac{2\theta}{\rho_1 + \theta(1 - r_1)} \int_{-\infty}^\infty \frac{1}{p_k(T)\,q_k(T)} \int_T^\infty p_k(\tau_1)\,q_k(\tau_1) \times e^{-(\rho_2+\theta)(\tau_1 - T)}\,d\tau_1 \int_T^\infty \left(\frac{1}{p_k(\tau_2)\,q_k(\tau_2)} \times \int_{\tau_2}^\infty p_k(\tau_3)\,q_k(\tau_3)\,e^{-(\rho_2+\theta)(\tau_3-\tau_2)}\,d\tau_3 \right) e^{-(\rho_1+\rho_2-\chi+\theta)(\tau_2 - T)} \times p_k(\tau_2)\,q_k(\tau_2)\,d\tau_2\,dT. \quad (68)$$

Gathering terms, we have

$$\frac{2\theta}{\rho_1 + \theta(1 - r_1)} \int_{-\infty}^\infty \int_T^\infty \int_T^\infty \int_{\tau_2}^\infty \times \frac{p_k(\tau_1)\,q_k(\tau_1)}{p_k(T)\,q_k(T)} \times e^{-(\rho_2+\theta)(\tau_1+\tau_3-2T)} e^{-(\rho_1-\chi)(\tau_2 - T)}\,d\tau_3\,d\tau_2\,d\tau_1\,dT. \quad (69)$$

Into the above, we substitute for p_k and q_k using (4) and make a change of variables to $z = e^T$ and $x_i = e^{\tau_i}$ to obtain

$$\frac{2\theta}{\rho_1 + \theta(1 - r_1)} \int_0^\infty \int_z^\infty \int_z^\infty \int_{x_2}^\infty \frac{(1 + z)^2}{x_2 z^2 (1 + x_1)^2} \times (1 + x_3)^2 \times \left(\frac{x_1 x_3}{z^2} \right)^{-(\rho_2+\theta)} \left(\frac{x_2}{z} \right)^{-(\rho_1-\chi)} dx_3\,dx_2\,dx_1\,dz. \quad (70)$$

Reversing the order of integration, the above becomes

$$\frac{2\theta}{\rho_1 + \theta(1 - r_1)} \left(\int\int_{x_1 > x_3} \int_0^{x_3} \int_z^{x_3} \frac{(1 + z)^2}{x_2 z^2 (1 + x_1)^2} \times (1 + x_3)^2 \times \left(\frac{x_1 x_3}{z^2} \right)^{-(\rho_2+\theta)} \left(\frac{x_2}{z} \right)^{-(\rho_1-\chi)} dx_2\,dz\,dx_3\,dx_1 + \int\int_{x_1 > x_3} \int_0^{x_3} \int_z^{x_1} \frac{(1 + z)^2}{x_2 z^2 (1 + x_1)^2 (1 + x_3)^2} \times \left(\frac{x_1 x_3}{z^2} \right)^{-(\rho_2+\theta)} \left(\frac{x_2}{z} \right)^{-(\rho_1-\chi)} dx_2\,dz\,dx_3\,dx_1 \right).$$

At this point, the innermost integral may be evaluated, the remaining terms expanded in z, and integration with respect to z performed. Substitution of sw for x_3 and s/w for x_1 allows integration over s. Finally, substitution of f for w^2 and successive integration by parts (setting terms such as $f^{\rho_2+\theta}$ to u and the remainder to dv) provides a solution for the first part of the net effect of a substitution with a modifier:

$$\frac{2\theta}{C} \left(\frac{2A\,\psi'(A)}{B(4A^2 - 1)} - \frac{2A(A + B)(\psi'(A + B) + \psi'(A))}{B(2A + B)((2A + B)^2 - 1)} + \frac{1}{A(2A + 1)(1 + 2A + B)} + \frac{2}{(A + B)(2A + B)(1 + 2A + B)} \right), \quad (71)$$

where $A = \rho_2 + \theta$, $B = \rho_1 - \chi$, $C = \rho_1 + \theta(1 - r_1)$, and ψ' is the trigamma function. When recombination rates are large ($\rho_1, \rho_2 > 1$), the above is approximately equal to

$$\frac{2\theta}{A^2 C(A + B)}. \quad (72)$$

Second part of (66): By reversing the order of integration, we get

$$-\int_{-\infty}^\infty r_1 p_k(\tau)\,q_k(\tau)\,\delta\Delta \frac{1}{\rho_1 + \theta(1 - r_1)}\,d\tau. \quad (73)$$

Substituting (68) for $\delta\Delta$ and gathering terms, we obtain

$$\frac{r_1}{\rho_1 + \theta(1 - r_1)} \int_{-\infty}^\infty \int_T^\infty \int_{\tau_1}^\infty p_k(\tau_2)\,q_k(\tau_2) \times e^{(\rho_2+\rho_1-\chi+\theta)T - (\rho_1-\chi)\tau_1 - (\rho_2+\theta)\tau_2}\,d\tau_2\,d\tau_1\,dT.$$

Into the above, we substitute for p_k and q_k using (4) and make a change of variables to $z = e^T$ and $x_i = e^{\tau_i}$ to obtain

$$\frac{r_1}{\rho_1 + \theta(1 - r_1)} \int_0^\infty \int_z^\infty \int_{x_1}^\infty \frac{1}{z x_1 (1 + x_2)^2} \times z^{(\rho_2+\rho_1-\chi+\theta)} x_1^{-(\rho_1-\chi)} x_2^{-(\rho_2+\theta)}\,dx_2\,dx_1\,dz.$$

By changing the order of integration to $dx_1\,dz\,dx_2$ and integrating, we find that the second part of the net effect equals

$$\frac{r_1}{AC(A + B)}, \quad (74)$$

where again $A = \rho_2 + \theta$, $B = \rho_1 - \chi$, and $C = \rho_1 + \theta(1 - r_1)$.

Summing both parts together, we get that the net

effect of a modifier on the fixation probability is equal to

$$\frac{r_1}{AC(A+B)} + \frac{2\theta}{C}\left(\frac{2A\,\psi'(A)}{B(4A^2-1)} - \frac{2A(A+B)(\psi'(A+B)+\psi'(A))}{B(2A+B)((2A+B)^2-1)} + \frac{1}{A(2A+1)(1+2A+B)} + \frac{1}{(A+B)(2A+B)(1+2A+B)}\right), \quad (75)$$

which reduces to

$$\frac{2\frac{\theta}{A} + r_1}{AC(A+B)}, \quad (76)$$

when recombination rates are large ($\rho_1, \rho_2 > 1$). When both r_1 and r_2 are much greater than s and S, (76) simplifies further to $S^3/(r_2(r_1 + r_2 - r_1 r_2))$.

To assess the approximations, the net effect of the modifier on fixation probability was calculated using the exact iteration for $P_{mk}[t]$ given by (1) and its analogues. Equations 58 for tight linkage and (75) and (76) were then compared to exact numerical solutions for the net effect in Table 3. A good approximation is provided by (75) when ρ_1, ρ_2, and θ are all <1, and by (76) when ρ_1 and ρ_2, or θ are >1.

APPENDIX C
HITCHHIKING AT THE MODIFIER LOCUS

Consider a new beneficial allele (j_1) that arises at random within a population in a single copy (at frequency of $p_j = 1/(2N)$). For now, other selected loci are assumed to be fixed and we focus on changes at the modifier locus due to hitchhiking. j_1 will arise in coupling with the modifier allele m_1 with probability p_m, leading to an initial linkage disequilibrium of $D = q_m p_j$. Otherwise, j_1 will arise with m_0 with probability $q_m = 1 - p_m$ leading to an initial disequilibrium of $D = -p_m p_j$. Using the standard two-locus recursions, we can write the change in frequency at the neutral modifier locus (M) as a function of the change in frequency at the selected locus (J) and the disequilibrium between M and J: $\delta p_m = (D/(p_j q_j))\,\delta p_j$. With an estimate of this quantity, we will be able to integrate δp_m over the entire sweep of the j_1 allele to determine the overall change in frequency at the modifier locus.

If selection is weak and stochastic effects are ignored, the dynamics of the beneficial mutation j_1 can be described by the logistic equation:

$$p_j = \frac{e^{st}}{(1+e^{st})},$$

where again t is measured from the mid-point of the sweep and $p_j = 1/(2N)$ at $t = t_0$. There will, however, be a large stochastic component to the change in gene frequency, especially when j_1 is rare. If we were to repeat the process a number of times, we would see that some mutations would be lost immediately, others would remain in the population for some time but then be lost, and only rarely would the new mutation fix. Of course, those processes in which, by chance, the mutation happens to rise faster than expected in the first few generations (over-sampling of the mutation) will be more likely to lead to fixation of the beneficial mutation. This leads to an apparent acceleration of the logistic equation for p_j during the first few generations when looking only at those alleles that do fix (MAYNARD-SMITH and HAIGH 1974; BARTON 1995b). Solution of the associated diffusion equation shows that the trajectory of the beneficial mutation is given by

$$p_j = \frac{e^{s(t+\tau)}}{(1+e^{s(t+\tau)})} \quad (77)$$

(BARTON 1995b and unpublished results), where τ measures the apparent acceleration in the first few generations (see below).

When the beneficial allele first arises with m_1, the initial disequilibrium equals $D/(p_j q_j) \approx q_m$. $D/(p_j q_j)$ measures the difference between the allele frequency of j_1 on m_1 chromosomes and its frequency on m_0 chromosomes. Recombination acts to equilibrate the allele frequencies on all backgrounds and $D/(p_j q_j)$ decays at a rate r_1, so that

$$E\left[\frac{D}{p_j q_j}\right] = q_m e^{-r_1 \hat{T}}, \quad (78)$$

where $\hat{T}$ measures time since j_1 was introduced ($\hat{T} = t - t_0$). Since the M locus is neutral, the disequilibrium trajectories that lead to fixation are not different from those that lead to loss; consequently, no acceleration term enters into this equation. We wish next to replace time, $\hat{T}$, in (78) with equations for allele frequencies. To begin, note that (77) may be written as

$$\frac{p_j}{q_j} = e^{s(t+\tau)} = e^{s(\hat{T}+t_0+\tau)} = e^{s(\hat{T}+\tau)}\frac{1}{2N}. \quad (79)$$

Hence,

$$e^{-r_1\hat{T}} \approx \left(\frac{p_j}{q_j}\right)^{-r_1/s}\left(\frac{e^{s\tau}}{2N}\right)^{r_1/s}.$$

One can show using the diffusion approximation that $z = 2s\,e^{s\tau}$ has an exponential distribution with parameter $\lambda = 1$ (N. H. BARTON, unpublished results). Therefore,

$$E[(e^{s\tau})^{r_1/s}] = E\left[\left(\frac{z}{2s}\right)^{r_1/s}\right] = (2s)^{-r_1/s}\Gamma\left(1+\frac{r_1}{s}\right).$$

Hitch of a modifier with tight linkage: We can now

estimate the total change in p_m when the beneficial allele arises with the modifier allele m_1 ($\Delta_1 p_m$) by integrating $\delta p_m = D/(p_j q_j)\,\delta p_j$ over the entire sweep:

$$\Delta_1 p_m = \int_{1/2N}^{1} \frac{D}{p_j q_j}\, dp_j$$

$$\approx q_m (4Ns)^{-r_1/s} \Gamma\left(1 + \frac{r_1}{s}\right) \int_0^1 \left(\frac{p_j}{q_j}\right)^{-r_1/s} dp_j$$

$$= q_m (4Ns)^{-r_1/s} \Gamma\left(1 + \frac{r_1}{s}\right)\Gamma\left(1 + \frac{r_1}{s}\right)\Gamma\left(1 - \frac{r_1}{s}\right), \quad (80)$$

which uses the fact that the integral describes a beta function that can be evaluated in terms of Γ functions as long as $r_1 < s$ (ABROMOWITZ and STEGUN 1970, p. 258). For small rates of recombination ($r_1 \ll s$) this becomes

$$\Delta_1 p_m \approx q_m\left(1 - \frac{r_1}{s}(\log(4Ns) + \gamma)\right), \quad (81)$$

where γ is Euler's constant. Similarly, if the mutation arises with allele m_0, the frequency of the m_1 allele changes by an amount

$$\Delta_0 p_m \approx -p_m (4Ns)^{-r_1/s} \Gamma\left(1 + \frac{r_1}{s}\right)^2 \Gamma\left(1 - \frac{r_1}{s}\right). \quad (82)$$

Equations 80 and 82 describe the change in frequency of the modifier locus given that j_1 is initially linked with a specific allele at locus M and given that j_1 does fix.

Taking into account the fact that the beneficial allele j_1 has a chance p_m of arising with m_1 in which case its fixation probability is F_1 and a chance q_m of arising with m_0 in which case its fixation probability is F_0, the unconditional expectation for the change in the modifier is

$$\overline{\Delta p_m} = p_m F_1 \Delta_1 p_m + q_m F_0 \Delta_0 p_m. \quad (83)$$

When selection acts only at locus J, the probability of fixation of j_1 is $F_1 = F_0 \approx 2s$, regardless of which allele is carried at the modifier locus. Therefore, from (80) and (82), $\overline{\Delta p_m}$ equals 0 as expected; hitchhiking does not, on average, change the allele frequencies at a neutral linked locus.

When more than one locus is under selection, however, the fixation probabilities differ from $2s$ as described by (11) and (12). Since $\delta\Pi$ is always positive under the conditions explored in this paper, the fixation of a beneficial mutation is more likely if it first appears linked to a modifier allele that increases recombination (here, m_1); that is, $F_1 > F_0$. Using (11) and (12), we can rewrite (83) as

$$\overline{\Delta p_m} = \frac{2s\,\delta r_2\,\delta\Pi}{S} p_m q_m (4Ns)^{-r_1/s} \times \Gamma\left(1 + \frac{r_1}{s}\right)^2 \Gamma\left(1 - \frac{r_1}{s}\right). \quad (84)$$

The strength of selection acting indirectly on the modifier locus is therefore,

$$s_m = \frac{\Delta p_m}{p_m q_m} = \frac{2s\,\delta r_2\,\delta\Pi}{S} (4Ns)^{-r_1/s} \Gamma\left(1 + \frac{r_1}{s}\right)^2 \Gamma\left(1 - \frac{r_1}{s}\right). \quad (85)$$

Note that this selection coefficient is not measured per generation but per mutation that arises when there is directional selection at another locus.

Hitch of a modifier with loose linkage: With loose linkage the situation is much simpler. If r_1 is large relative to s, then the disequilibrium decays rapidly while j_1 is still rare. Therefore we can assume that $D/(p_j q_j) \approx D/p_j$. To obtain an estimate of the change in frequency at the modifier locus, we note that the standard two-locus recursions can be used to show that the change in frequency of a neutral linked allele (p_m) per generation depends only on the disequilibrium between the two loci: $\Delta p_m = Ds$. We therefore focus on obtaining an estimate for D. For clarity, we will write D as $D(\hat{T})$, to emphasize the dependence of D on time. As long as the disequilibrium decays while j_1 is still rare, $D(\hat{T})/p_j$ will decrease at rate r_1. Furthermore, we can use (79) to note that $p_j \approx e^{s(\hat{T}+\tau)}/2N$ conditional upon the fixation of j_1. To obtain an estimate of $D(\hat{T})$, first note that

$$E\left[\frac{D(\hat{T})}{p_j}\right] \approx E\left[\frac{D(\hat{T})}{e^{s(\hat{T}+\tau)}/2N}\right] \approx \frac{E[D(\hat{T})]}{e^{s\hat{T}}(1/2s)(1/2N)}, \quad (86)$$

using the fact that $z = 2s\,e^{s\tau}$ has an exponential distribution with parameter $\lambda = 1$ independent of the value of $D(\hat{T})$. Rearranging,

$$E[D(\hat{T})] = e^{s\hat{T}} \frac{1}{2s}\frac{1}{2N} E\left[\frac{D(\hat{T})}{p_j(\hat{T})}\right]$$

$$\approx e^{s\hat{T}} \frac{1}{4sN} e^{-r_1\hat{T}} \frac{D(0)}{p_j(0)} \approx e^{(s-r_1)\hat{T}} \frac{1}{4sN}\frac{D(0)}{p_j(0)}. \quad (87)$$

When j_1 arises in coupling with m_1, $D(0)/p_j(0) \approx q_m$. In this case, the expected hitch of the loosely linked modifier allele is

$$\Delta_1 p_m = \int_0^\infty s\,D(\hat{T})\,d\hat{T} \approx \int_0^\infty s e^{(s-r_1)\hat{T}} \frac{1}{4sN} q_m = \frac{q_m}{4N(r_1 - s)}, \quad (88)$$

When j_1 arises in coupling with m_0, $D(0)/p_j(0) \approx p_m$ and the change in frequency of the m_1 allele is

$$\Delta_0 p_m = -\frac{p_m}{4N(r_1 - s)}. \quad (89)$$

When only locus J is under selection, the probability of fixation of j_1 is $F_1 = F_0 \approx 2s$. In this case, using (83), the unconditional expectation for the change in frequency of m_1 is again 0. With selection at more than one locus, we must use (11) and (12) for F_1 and F_0. With loose linkage, therefore, we have that the expected change in frequency at the modifier locus is

$$\overline{\Delta p_m} = p_m F_1 \frac{q_m}{4N(r_1 - s)} + q_m F_0 \frac{-p_m}{4N(r_1 - s)}$$

$$= p_m q_m \frac{1}{4N(r_1 - s)} \frac{2s\,\delta r_2\,\delta\Pi}{S}. \quad (90)$$

Simulation check: A two-locus model with selection acting on only one locus, J, was simulated to check (80) and (88). The population (size $2N = 10^6$) was begun with a single copy of allele j_1 initially linked with m_1 at the second, neutral locus. The frequency of m_1 was set to $p_m = 1/2$ at the beginning of the simulation and was monitored during the sweep of j_1. Among those processes that fixed, Δp_m was determined and compared to (80) and (88) for a variety of values of the recombination rate as shown in Figure 12. The equations derived here provide better estimates of the change in allele frequency than do the deterministic estimates provided by MAYNARD SMITH and HAIGH (1974, their Equation 14) or by STEPHAN *et al.* (1992, their Equation 17) as shown in Figure 12 (dotted and lower dashed lines, respectively). Our approximations perform better than these deterministic estimates because they take into account the initial stochastic fluctuations that occur when the beneficial mutation first arises. Specifically, we incorporate the apparent acceleration in the spread of a beneficial allele that occurs among those processes that successfully lead to the fixation of the beneficial allele. In a previous study, STEPHAN *et al.* (1992) used a diffusion approach to incorporate stochastic fluctuations into estimates of the effect of hitchhiking. These authors focused, however, on estimating the expected change in heterozygosity (Equation 14c). Since heterozygosity is not a linear function of allele frequencies, the expected change in heterozygosity does not provide a good estimate of the expected change in allele frequency (see upper dashed lines in Figure 12). We therefore use (80) and (88) to describe the expected change in frequency of an allele due to hitchhiking with a new beneficial mutation.

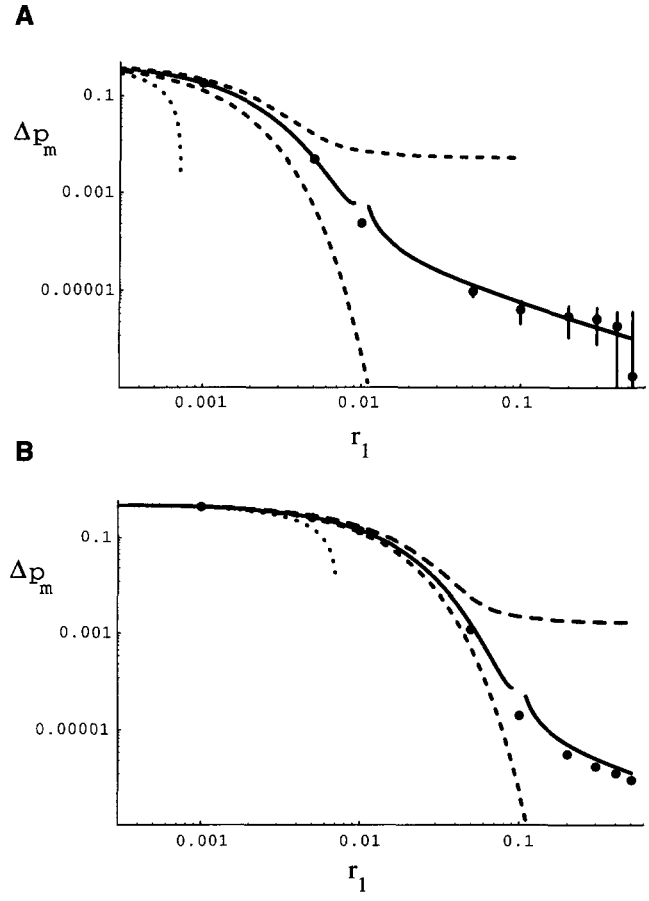

FIGURE 12.—The change in allele frequency due to hitchhiking. The change in frequency, Δp_m, of a neutral allele that is initially associated with a beneficial allele at a linked locus is shown, given that the beneficial allele does fix within the population. Simulation results are given by the dots for various values of r_1 between the neutral and selected loci, with parameters $2N = 10^6$, $p_m(0) = 1/2$, $p_j(0) = 1/(2N)$, and with 10^6 replications per dot. The vertical lines denote ±1.96 SE for each estimate. —, (80) for small r_1 and (88) for large r_1; · · ·, estimates based on (14) from MAYNARD SMITH and HAIGH (1974); ---, estimates based on (14c) (bottom dashed line) and (17) (upper dashed line) from STEPHAN *et al.* (1992). (A) $s = 0.01$. (B) $s = 0.1$.