The Hill–Robertson Effect and the Evolution of Recombination

Denis Roze[1] and Nick H. Barton

Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

ABSTRACT

In finite populations, genetic drift generates interference between selected loci, causing advantageous alleles to be found more often on different chromosomes than on the same chromosome, which reduces the rate of adaptation. This "Hill–Robertson effect" generates indirect selection to increase recombination rates. We present a new method to quantify the strength of this selection. Our model represents a new beneficial allele (A) entering a population as a single copy, while another beneficial allele (B) is sweeping at another locus. A third locus affects the recombination rate between selected loci. Using a branching process model, we calculate the probability distribution of the number of copies of A on the different genetic backgrounds, after it is established but while it is still rare. Then, we use a deterministic model to express the change in frequency of the recombination modifier, due to hitchhiking, as A goes to fixation. We show that this method can give good estimates of selection for recombination. Moreover, it shows that recombination is selected through two different effects: it increases the fixation probability of new alleles, and it accelerates selective sweeps. The relative importance of these two effects depends on the relative times of occurrence of the beneficial alleles.

O NE of the first general hypotheses proposed to explain the maintenance of sex and recombination states that sex increases the rate of adaptation, by allowing beneficial mutations initially present on different chromosomes to be combined onto the same chromosome (MORGAN 1913; FISHER 1930; MULLER 1932). The first analytical treatment of the question was done in the 1960s and led to some debate: using a simple calculation, CROW and KIMURA (1965) showed that the rate of incorporation of beneficial mutations is indeed higher in sexual than in asexual populations, whereas a deterministic model made by MAYNARD SMITH (1968) showed no difference between sexual and asexual populations. FELSENSTEIN (1974) pointed out that this disagreement stemmed from the different assumptions made by these models: Crow and Kimura represented a finite population in which advantageous alleles first arise as single copies, while Maynard Smith considered the case of an infinite population, in which deleterious alleles at two loci are maintained at mutation–selection equilibrium and become advantageous after a change in the environment. Because the model assumes no epistasis, the two loci always stay in linkage equilibrium, and recombination has no effect on the change in genotype frequencies. By contrast, in a finite population such as the one modeled by Crow and Kimura, selection and random drift combine to generate some linkage disequilibrium even in the absence of epistasis. For example, consider

the case where a new beneficial allele enters the population while another beneficial allele at a second locus is already present. The new mutation can occur either on the good background at the second locus, in which case the linkage disequilibrium (LD) is positive, or on the bad background, in which case the linkage disequilibrium is negative. A simple calculation shows that, on average, this initial linkage disequilibrium is zero. However, when both advantageous alleles arise in coupling (LD > 0) they tend to increase in frequency rapidly, while when they are in repulsion (LD < 0) they compete against each other, which slows their rate of increase. Because situations in which LD < 0 tend to last longer, the initial variance in linkage disequilibrium generates, on average, negative linkage disequilibrium. The same "Hill–Robertson effect" occurs when the variance in LD is generated by drift due to finite population size in a single population (HILL and ROBERTSON 1966; FELSENSTEIN 1974) or in an arbitrary number of interconnected populations (MARTIN *et al.* 2006). The negative linkage disequilibrium generated by the Hill–Robertson effect reduces the rate of incorporation of advantageous alleles and limits the speed of adaptation. BARTON (1995b) quantified this effect by calculating the reduction in fixation probability of an advantageous allele when selection occurs simultaneously at another linked locus.

Increasing the recombination rate between loci decreases the strength of the Hill–Robertson effect and thus increases the rate of fixation of beneficial alleles. Simulation and analytical models have shown that this generates indirect selection on modifier genes that increase recombination rates (FELSENSTEIN and YOKOHAMA

---

[1]*Corresponding author:* Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, King's Bldgs., W. Mains Rd., Edinburgh EH9 3JT, United Kingdom. E-mail: denis.roze@ed.ac.uk

1976; OTTO and BARTON 1997, 2001; ILES *et al.* 2003; BARTON and OTTO 2005; MARTIN *et al.* 2006). Recombination modifiers are not mere theoretical constructs, but have been identified in a number of experimental studies. In particular, selection experiments have shown that it is possible to artificially increase recombination within genomes, either by direct selection on recombination rates or as a by-product of selection for another trait (*e.g.*, KIDWELL 1972; CHARLESWORTH and CHARLESWORTH 1985a,b; KOROL and ILIADI 1994; KOROL *et al.* 1994; and references in OTTO and BARTON 2001). In some cases, genetic factors affecting recombination rates have been mapped, revealing a variety of effects: some factors have a local effect, increasing recombination only over a small region of the genome (which can be located on the same or on a different chromosome), while others affect the overall recombination rate of whole chromosomes. Recombination rates are also known to evolve in nature: for example, TRUE *et al.* (1996) found that the genetic map of the island endemic *Drosophila mauritiana* is 1.8 times longer than the map of its sister species *D. melanogaster*, with a higher frequency of crossovers along each chromosome. Rapid adaptation to a new environment may have selected for a higher recombination rate, but a broader survey is needed to establish whether there is a correlation between adaptation and recombination rate.

The first analytical study on the evolution of recombination due to the Hill–Robertson effect was by OTTO and BARTON (1997). Their model represents two loci under directional selection: at the first locus, a new beneficial allele arises as a single copy, while at the second locus another beneficial allele is already established and sweeps through the population; finally, a third locus affects the recombination rate between the selected loci. Because recombination decreases interference between the selected loci, the new beneficial allele has a higher fixation probability when it occurs on the high-recombination background. The frequency of the modifier allele coding for higher recombination is then expected to increase by hitchhiking, on average, as the new beneficial allele goes to fixation. Otto and Barton's analysis proceeds in two steps: first, they derive the probability of fixation of the new advantageous allele, arising on each of the four possible backgrounds at the two other loci; then, they use a deterministic calculation to obtain the change in frequency of the recombination modifier, given that the new beneficial allele reaches fixation. This approach, however, has several limitations. In particular, although interference between selected loci due to the Hill–Robertson effect is taken into account in the calculation of fixation probabilities, it does not enter the deterministic calculation of the hitch given by the new beneficial allele to the modifier, given that it fixes. This last calculation indeed assumes linkage equilibrium between the se-

lected loci. There are, however, three possible sources of LD, which, as discussed by Otto and Barton, are probably the cause of the quantitative differences they observe between analytical and simulation results. The first is the initial disequilibrium generated by the mutation event producing the new beneficial allele; although this disequilibrium is zero on average, it is sometimes positive (when the new allele arises in coupling with the other beneficial allele) and sometimes negative (when it arises in repulsion), and the effect of this disequilibrium on the change in frequency of the modifier does not cancel out. This initial disequilibrium is taken into account in some of the numerical calculations of Otto and Barton (dotted lines in Figures 4–9 of OTTO and BARTON 1997), but this still leaves important departures from simulation results in several cases. The second source of disequilibrium is random drift, which can have strong effects while the new allele is still rare. The effect of this disequilibrium is analyzed in a more recent model (BARTON and OTTO 2005), by considering small fluctuations around the deterministic increase of the new allele. This analysis, however, does not apply while the allele is present in just a few copies (in which case drift can generate important departures from the deterministic trajectory), which is just when the effect on the modifier is strongest. Finally, a third source of disequilibrium comes from the fact that when the new beneficial allele happens to reach fixation, it is more likely to be in coupling with the beneficial allele at the other selected locus. Therefore, the disequilibrium on the paths that lead to fixation of the new allele tends to be more positive than the disequilibrium observed on the paths leading to loss of the allele.

In this article, we show that these restrictions can be overcome by finding the distribution of numbers of the new mutation across the four backgrounds, after it has become established but while it is still rare. Since this initial stage involves the independent replication of rare mutants, it can be described by a branching process; the solution can then be taken forward deterministically, to find the net expected increase of the modifier allele. This assumes that population size is sufficiently large that the effects of drift can be neglected once alleles are established within the population. As we will see, this method involves solving numerically a set of differential equations; therefore, one has to resort to a computer to evaluate the strength of selection on the modifier, as with simulation methods. However, our method has two main advantages over simulations: first, it can stand in place of the millions of replicate simulations needed to get accurate estimates of the change in frequency of the modifier; second (and more importantly), it provides us with an understanding of the scaling relationships among the parameters. We will see that our method provides accurate predictions over a wide range of parameter values. It also captures two different effects of

recombination modifiers: they affect the fixation probability of new advantageous alleles and also the speed of selective sweeps. These two effects both generate indirect selection at the modifier locus; we will see that their relative importance depends on the relative times of occurrence of beneficial alleles. Finally, we use our method to estimate the overall strength of selection on a modifier affecting recombination over the whole genome, when recurrent beneficial mutations occur throughout the genome. This shows that the rate of adaptive substitutions has to be relatively high for interactions between selected alleles to generate a substantial force on recombination modifiers. Although the estimated rates of adaptive amino acid substitutions for the Drosophila and the human lineages do not seem compatible with such high values, it is still possible that most adaptive changes occurred in relatively short periods, during which indirect selection on recombination may have been important. Despite this requirement for a high substitution rate, stochastic forces may still appear as a more likely candidate to explain the widespread occurrence of recombination than deterministic forces generated by epistasis among loci. Indeed, epistasis can select for increased recombination rates only when it is negative and not too variable among loci (BARTON 1995a; OTTO and FELDMAN 1997), whereas available data do not show any clear tendency for epistasis to follow this pattern. Finally, we will see that our model does not give accurate results for all possible parameter values and initial conditions: it best describes the case where a new, relatively weakly selected allele occurs when a more strongly selected allele is already sweeping through the population. We have also worked on another method that better describes the case where a relatively strongly selected allele occurs while a more weakly selected one is already established and the case where two beneficial alleles occur at about the same time (N. H. BARTON and D. ROZE, unpublished results). In these cases, recombination has little influence on fixation probabilities, but still has an effect in accelerating selective sweeps. Finally, the method of BARTON and OTTO (2005) deals with the effects of small fluctuations due to drift, once alleles are abundant. Together, these methods thus give us a complete understanding of the different regimes under which recombination can be favored.

GENERAL METHOD

The general setting is the same as in OTTO and BARTON (1997). We consider a haploid population of size N; mating is random, and generations are discrete and nonoverlapping. We assume that selection occurs at two loci, favoring allele A over allele a at the first locus and allele B over allele b at the second locus. We call s_A and s_B the selective advantages of A and B and assume no epistasis, so that the fitnesses of the different genotypes are given by

$$W_{ab} = 1 \qquad W_{Ab} = 1 + s_A$$
$$W_{aB} = 1 + s_B \qquad W_{AB} = (1 + s_A)(1 + s_B). \qquad (1)$$

We assume that $N \gg 1/s_A$, and $N \gg 1/s_B$, so that the fate of alleles A and B is decided while they are still rare. We then assume that a third locus, with two alleles M and m, affects the recombination rate between the two selected loci (recombination modifier). The recombination rate between loci (A, a) and (B, b) equals $r_{AB}(1 - \delta r)$, r_{AB}, and $r_{AB}(1 + \delta r)$ in mm, Mm, and MM individuals, respectively; δr thus measures the modifier effect, and for simplicity we suppose that this effect is additive. The recombination rate between loci (M, m) and (A, a) is denoted r_{MA}. We present the model in the case where loci are in order (M, m)–(A, a)–(B, b); the method applies more generally, however, and we also give results for different orderings of the three loci.

We call t the time at which allele A enters the population (as a single copy). We suppose that, at this time, allele B is already established (although it can still be rare). We see later that this assumption is not crucial and that in many cases we can still obtain accurate results even if B is not established at t. We also assume that allele M is frequent enough that stochastic fluctuations can be neglected at the (M, m) locus and suppose that, at time t, the (M, m) and (B, b) loci are in linkage equilibrium (thus, we neglect any associations between M and B that may have been generated by previous sweeps at yet other loci).

We decompose the spread of allele A into two phases: during the first phase, between times t and t^*, A either is lost from the population by drift or becomes established (i.e., reaches a high enough number of copies so that it will almost certainly fix). In the following we assume that A is still rare in the population at t^*, so that between t and t^*, the replication of each copy of A does not depend on the presence of other A alleles in the population, in which case the change in number of A's can be described by a branching process. After t^* we assume that A, if not lost, increases in frequency on a deterministic trajectory. As we said before, allele B at the other selected locus is supposed to be already established at t, and we assume that the frequency of B stays on a deterministic trajectory all the time. We suppose that t^* is sufficiently far that B will have reached fixation at t^*. We assume that the frequency of the modifier allele M is not too small, so that M changes in frequency solely due to its effect on the spread of allele A: when M helps A to recombine onto the good B background, it then tends to be carried along during the spread of A after time t^*, by hitchhiking. Before t^*, the effect of A on the frequency of the modifier is very small (because A remains at low frequency). Thus, we assume that the frequency of M remains constant between t and t^*.

In summary, the only stochastic process in our model concerns the change in number of copies of A between

times t and $t*$. During this time, the frequency of M does not change, while B increases in frequency on a deterministic trajectory, until fixation. After $t*$, B is fixed, A increases on a deterministic trajectory if it is still present, and M changes in frequency due to the linkage disequilibrium established with A during the first phase. Note that these different assumptions all stem from the assumption that population size is large and that M and B are common enough at t that the effects of drift at these loci can be neglected.

In APPENDIX A, we show that the expected total change in frequency of the recombination modifier after time $t*$ is given by

$$E[\Delta p_M] \approx N^{-\rho}\Gamma(1-\rho)\Gamma(1+\rho)E[S^{\rho-1}D], \qquad (2)$$

where $\rho = r_{MA}/s_A$, S is the total number of A alleles at $t*$, and D is the linkage disequilibrium between loci (M, m) and (A, a) at $t*$, multiplied by population size N. To express the expected change in frequency of the modifier, we thus need to calculate the expectation of $S^{\rho-1}D$ over the probability distribution of the number of copies of A on the different possible backgrounds at $t*$. We now present a method to derive this distribution.

In the following, we use index j to denote a possible genetic background on which allele A may be found (in our three-locus model, the possible backgrounds are thus mb, mB, Mb, and MB). As we said before, we assume that A is present as a single copy at time t; however, let us assume for the moment that the initial number of copies of A is arbitrary and call $n_{j;0}$ the number of copies of A on background j at t. We denote by $\mathbf{n}_0$ the vector of all $n_{j;0}$. We then call n_j the number of copies of A on background j at time $t*$ and $\mathbf{n}$ the vector of all n_j. Finally, $\phi(\mathbf{n})$ is the probability that the number of copies of A on the different backgrounds at $t*$ is equal to $\mathbf{n}$. Although it is difficult to derive the distribution ϕ directly, we show below that it is possible to obtain its Fourier transform ψ, defined using the vector of dummy variables $\boldsymbol{\omega}$,

$$\psi(\boldsymbol{\omega}) = \int_{\mathbf{n}} \phi(\mathbf{n}) e^{-\iota \sum_j n_j \omega_j} d\mathbf{n}, \qquad (3)$$

where ι is the imaginary number $\sqrt{-1}$, and where the integral is a multiple integral over all possible values of $\mathbf{n}$. Note that although $\mathbf{n}$ takes only discrete values, we treat it here as a continuous variable (since we integrate over $\mathbf{n}$, rather than summing); this will later allow us to use a diffusion approximation. The Fourier transform ψ is thus a function of variables ω_j, and $\boldsymbol{\omega}$ is the vector of all ω_j's (which has the same dimension as $\mathbf{n}$). Considering ψ rather than ϕ proves useful, because we can make use of the convolution theorem, which states that the Fourier transform of the distribution of a sum of independent random variables is equal to the product of the Fourier transforms of the individual distributions of these variables. Because we assume that copies of allele A replicate independently between t and $t*$, the distributions

of the number of copies originating from each initial A allele are independent. Since $\mathbf{n}$ is the sum of these numbers, ψ must take the form

$$\psi(\boldsymbol{\omega}) = \prod_j [\psi_j(\boldsymbol{\omega})]^{n_{j;0}}, \qquad (4)$$

where $\psi_j(\boldsymbol{\omega})$ is the Fourier transform of the distribution of the number of copies of A at $t*$ originating from a single copy, present on background j at t (the functions ψ_j thus do not depend on the $n_{j;0}$'s). Equation 4 can also be written

$$\psi(\boldsymbol{\omega}) = e^{-\sum_j P_j n_{j;0}}, \qquad (5)$$

where each P_j equals $-\ln[\psi_j(\boldsymbol{\omega})]$. Note that instead of the Fourier transform, we could also use the Laplace transform or the generating function of ϕ, which also satisfies the convolution theorem. However, in the following we need to invert the transform, and for this the Fourier transform proves more convenient.

When $t = t*$, the distribution ϕ is reduced to a single point at $\mathbf{n} = \mathbf{n}_0$. The Fourier transform of such a distribution is given by $\psi(\boldsymbol{\omega}) = \exp(-\iota \sum_j n_{j;0}\omega_j)$, giving at $t = t*$

$$P_j = \iota\omega_j \qquad (6)$$

for all j. We then use a diffusion approximation to express the P_j's as a function of t, for a given value of $t*$. A backward diffusion on ψ is given by

$$\frac{\partial\psi}{\partial t} = -\sum_j \left[\left(\sum_k \sigma_{jk} n_{k;0} \right) \frac{\partial\psi}{\partial n_{j;0}} + \frac{n_{j;0}}{2} \frac{\partial^2\psi}{\partial n_{j;0}^2} \right]. \qquad (7)$$

The term $\sum_k \sigma_{jk} n_{k;0}$ in Equation 7 represents the expected increase in n_j at t, assuming that all $n_{j;0}$ are small relative to N (the term σ_{jk} represents the expected change in n_j accounting for possible movement from background k and can be expressed in terms of recombination rates and selection coefficients). Assuming a Poisson number of offspring per parent, the variance in the change of n_j at t is approximately $n_{j;0}$ (for $n_{j;0} \ll N$), while the covariance between the changes in n_j and n_k is negligible for all $j \neq k$. Substituting Equation 5 into Equation 7 then yields

$$\frac{\partial P_j}{\partial t} = -\left(\sum_k \sigma_{kj} P_k \right) + \frac{P_j^2}{2} \qquad (8)$$

(*e.g.*, HARRIS 1963, chap. V.15). Equation 8, together with the boundary condition (6), can be used to express the P_j's, and thus ψ, for arbitrary t. In general, Equation 8 will have to be solved numerically. One can then recover the distribution ϕ by taking an inverse Fourier transform or derive directly moments of ϕ from its Fourier transform ψ; this is illustrated in the next two-locus example. Because we use the diffusion approximation, we have to

assume that selection and recombination are weak (so that the n_j's change slowly over time). However, we will see that this is not a major restriction of our model. Finally, we note that Equation 8 is identical to the equation giving fixation probabilities of mutations occurring on different backgrounds in BARTON (1995b) and OTTO and BARTON (1997), except that boundary conditions differ. As we will see later, fixation probabilities can be recovered from the limit of the Fourier transform ψ as ω tends to infinity (this is because the distribution ϕ has a spike at zero, corresponding to the loss of the A allele, which is described by the limit of ψ for large ω). In the next section, we illustrate our method by considering the case of two selected loci (without a recombination modifier).

TWO-LOCUS CASE

We consider here the effect of selection at the (B, b) locus on the distribution of the number of copies of A at t^*, in the absence of a recombination modifier. Because we assume that B is fixed at t^*, allele A can be found only on background B at that time. The distribution ϕ is thus a function of a single variable (the number of A alleles at t^*, which will be on background B), which we simply denote n. The Fourier transform of $\phi(n)$ is given by

$$\psi(\omega) = \int_0^\infty \phi(n) e^{-\iota n \omega} dn. \tag{9}$$

We also know from the convolution theorem that ψ must take the form

$$\psi(\omega) = e^{-P_B n_{B;0} - P_b n_{b;0}} \tag{10}$$

(Equation 5), where P_B and P_b are functions of ω, and where $n_{B;0}$ and $n_{b;0}$ are the number of copies of A on the B and b backgrounds at time t, respectively. At $t = t^*$, the distribution ϕ is reduced to a single point at $n = n_{B;0}$, and its Fourier transform is thus given by $\psi(\omega) = e^{-\iota \omega n_{B;0}}$. This, with Equation 10, gives the boundary conditions

$$P_B = \iota \omega, \quad P_b = 0 \tag{11}$$

at $t = t^*$ (Equation 11 is a special case of Equation 6, constrained so that allele A is present only on the B background at t^*). From now on, we assume that A enters the population as a single copy at t. This copy will be on background B with probability p_B (which is the frequency of B at t) or on background b with probability p_b (which is the frequency of b). Equation 10 can thus be written

$$\psi(\omega) = p_B e^{-P_B(\omega)} + p_b e^{-P_b(\omega)}. \tag{12}$$

Indeed, $e^{-P_B(\omega)}$ is the Fourier transform $\psi(\omega)$ conditional on the initial copy of A being on background B (in which case $n_{B;0} = 1$ and $n_{b;0} = 0$ in Equation 10), while

$e^{-P_b(\omega)}$ is $\psi(\omega)$ conditional on the initial copy of A being on background b. Assuming that selection is relatively weak at the (B, b) locus, p_B and p_b are given by the logistic equations

$$p_B = \frac{e^{s_B t}}{1 + e^{s_B t}}, \quad p_b = \frac{1}{1 + e^{s_B t}}, \tag{13}$$

where $t = 0$ is at the midpoint of the sweep of allele B.

The σ-coefficients that appear in Equation 8 are easily obtained. The expected number of offspring copies produced by an A allele present on a b background at time t is $(1 + s_A)/(1 + s_B p_B) \approx 1 + s_A - s_B p_B$. Of these, a proportion $r_{AB} p_B$ moves to the B background, while a proportion $1 - r_{AB} p_B$ stays on the b background. The expected number of copies produced by an A allele present on a B background at t is $(1 + s_A)(1 + s_B)/(1 + s_B p_B) \approx 1 + s_A + s_B p_b$; of these, a proportion $r_{AB} p_b$ moves to the b background, while a proportion $1 - r_{AB} p_b$ stays on the B background. This gives (using Equation 8)

$$\frac{\partial P_b}{\partial t} = -(s_-)P_b + r_{AB} p_B (P_b - P_B) + \frac{P_b^2}{2}$$

$$\frac{\partial P_B}{\partial t} = -(s_+)P_B + r_{AB} p_b (P_B - P_b) + \frac{P_B^2}{2} \tag{14}$$

with

$$s_- = s_A - s_B p_B, \quad s_+ = s_A + s_B p_b. \tag{15}$$

These are the same as Equations 5a and 5b in BARTON (1995b). Equation 14, with boundary condition (11), can be solved numerically using the NDSolve routine of Mathematica (notebook available on request), giving ψ (Equation 12).

To recover the distribution $\phi(n)$ from the Fourier transform $\psi(\omega)$, it is useful to decompose ϕ into two parts: a distribution that is different from zero only when $n = 0$ (loss of allele A) and that is equal to $\phi(0)$ at this point [this distribution is thus given by $\phi(0)\delta(n)$, where δ is Dirac's "delta function"] and the rest of the distribution ϕ for $n > 0$ that we approximate by a continuous function on the range $n \geq 1$. The Fourier transform of the first part (which we call ψ_0) is a constant: $\psi_0(\omega) = \phi(0)$ for all ω [because the Fourier transform of $\delta(n)$ is 1], while the Fourier transform of the second part (which we call $\psi_{>0}$) tends to zero as ω tends to infinity (because the Fourier transform of a piecewise continuous function tends to zero as ω tends to infinity, e.g., LIGHTHILL 1958). Because ψ is the sum of ψ_0 and $\psi_{>0}$, its limit as ω tends to infinity is thus $\phi(0)$: therefore, we can obtain the probability that allele A is extinct at time t^* from the value of ψ for large ω. The rest of the distribution ϕ (for $n > 0$) can be recovered by inverting the Fourier transform $\psi_{>0}$. In APPENDIX B, we show that this can be done for each value of n by calculating a sum of n_{max} terms, where n_{max} is an arbitrary value such that the probability that allele A is present in
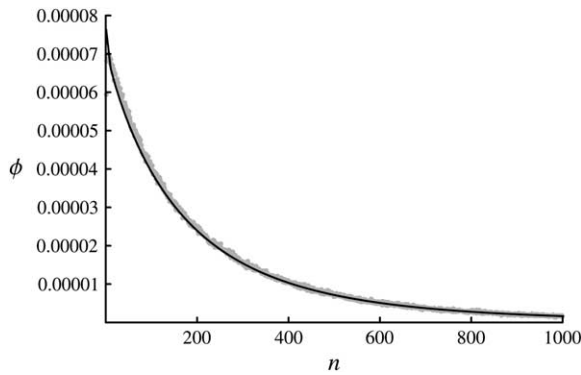
FIGURE 1.—Two-locus model: probability distribution of the number of copies of allele A at time t^* (n), in the absence of recombination modifier, from the analytical model (using Equation B6 with $n_{max} = 5000$, solid curve) and from simulations (shaded points). Parameter values are $s_A = 0.01$, $s_B = 0.1$, $r_{AB} = 0.02$, $t = -100$, $t^* = 100$, and $N = 10^6$. In the simulations, the distribution is obtained from 10^8 replications.

FIGURE 2.—Logarithm of the variance of the number of copies of allele A at time $t^* = 100$, as a function of the time t when allele A first appeared. Curves correspond to the model predictions, circles and squares to simulation results (variance observed among 10^7 replications). Solid curve, filled circles: two-locus model, $s_A = 0.01$, $s_B = 0.1$, $r_{AB} = 0.02$, $N = 10^6$. Dashed curve, squares: two-locus model, $s_A = 0.01$, $s_B = 0.1$, $r_{AB} = 0.04$, $N = 10^6$. Dotted curve: single selected locus (no interference), $s_A = 0.01$, $N = 10^6$.

a number of copies greater than n_{max} at time t^* can safely be ignored (remember that we assume that A is still rare at t^*). One obtains

$$\phi(n) = \frac{1}{n_{max}} \sum_{k=-n_{max}/2}^{n_{max}/2-1} [\psi(\omega_k) - \psi(\infty)] e^{\iota\omega_k n}, \quad \text{with } \omega_k = 2\pi \frac{k}{n_{max}} \tag{16}$$

and where $\psi(\infty)$ is the limit of ψ as ω tends to infinity (which can be determined by evaluating ψ for a very large value of ω or from Equation 6 in BARTON 1995b, giving fixation probabilities). The number of terms to be evaluated can then be reduced by using some symmetries of ψ (see APPENDIX B). In Figure 1 we compare this solution with simulation results, for $s_A = 0.01$, $s_B = 0.1$, $r_{AB} = 0.02$, $t = -100$, $t^* = 100$, and $N = 10^6$ (and where p_B is given by Equation 13). Note that the population size N does not appear in the derivation, but appears in the simulation program, which represents a Wright–Fisher model with multinomial sampling in each generation (our simulation program is written in C++ and is available on request). Figure 1 shows that Equation 16 accurately describes the probability distribution $\phi(n)$.

Moments of the distribution ϕ can be obtained directly from ψ, without having to invert the Fourier transform (in essence, the method is similar to Ohta and Kimura's method for deriving moments using the diffusion approximation, *e.g.*, OHTA and KIMURA 1969, 1971). For example, we show in APPENDIX B that the variance of ϕ is obtained by solving numerically a system of three differential equations, derived from Equation 14. Figure 2 shows the logarithm of the variance at time $t^* = 100$, as a function of the time t at which allele A first entered the population for $s_A = 0.01$, $s_B = 0.1$, $r_{AB} = 0.02$ (solid line), and $r_{AB} = 0.04$ (dashed line). The dotted line shows the logarithm of the variance in the absence
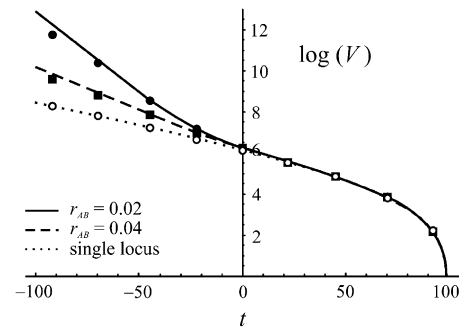
of interference (in the one-locus model), given by $e^{s_A \Delta t}(e^{s_A \Delta t} - 1)/s_A$, with $\Delta t = t^* - t$ (*e.g.*, FELLER 1951). Figure 2 illustrates the fact that increasing recombination between the selected loci decreases stochastic effects at the (A, a) locus (the variance is lower with $r_{AB} = 0.04$ than with $r_{AB} = 0.02$); this generates an indirect selective pressure to increase recombination rates, which we quantify in the next section.

CHANGE IN FREQUENCY AT THE MODIFIER LOCUS

We now introduce the recombination modifier locus (M, m). As we said before, because N is large we can assume that the frequency of allele M does not change between times t and t^*; we denote this frequency p_M, while p_m is the frequency of allele m. Furthermore, since we suppose that allele B is fixed at t^*, allele A can be found only on backgrounds MB and mB at t^*. Calling n_{MB} and n_{mB} the numbers of copies of A on these two backgrounds at t^*, the Fourier transform of the probability distribution $\phi(n_{MB}, n_{mB})$ is given by

$$\psi(\omega_{MB}, \omega_{mB}) = \iint \phi(n_{MB}, n_{mB}) e^{-\iota\omega_{MB} n_{MB} - \iota\omega_{mB} n_{mB}} dn_{MB}\, dn_{mB}. \tag{17}$$

From the convolution theorem, we also know that ψ must take the form

$$\psi(\omega_{MB}, \omega_{mB}) = e^{-\sum_j P_j n_{j;0}} \tag{18}$$

(Equation 5), where the sum is over all possible initial backgrounds $j = mb$, mB, Mb, and MB, and where the P_j's are functions of ω_{MB} and ω_{mB}. At $t = t^*$, we have boundary conditions

$$P_{MB} = \iota\omega_{MB}, \quad P_{mB} = \iota\omega_{mB}, \quad P_{Mb} = 0, \quad P_{mb} = 0. \tag{19}$$

From Equation 8, one then obtains the system

$$\frac{\partial P_{mb}}{\partial t} = -(s_-)P_{mb} - r_{AB}(1 - p_m\delta r)p_B(P_{mB} - P_{mb})$$
$$- r_{MA}p_M(P_{Mb} - P_{mb}) + \frac{P_{mb}^2}{2}$$

$$\frac{\partial P_{Mb}}{\partial t} = -(s_-)P_{Mb} - r_{AB}(1 + p_M\delta r)p_B(P_{MB} - P_{Mb})$$
$$- r_{MA}p_m(P_{mb} - P_{Mb}) + \frac{P_{Mb}^2}{2}$$

$$\frac{\partial P_{mB}}{\partial t} = -(s_+)P_{mB} - r_{AB}(1 - p_m\delta r)p_b(P_{mb} - P_{mB})$$
$$- r_{MA}p_M(P_{MB} - P_{mB}) + \frac{P_{mB}^2}{2}$$

$$\frac{\partial P_{MB}}{\partial t} = -(s_+)P_{MB} - r_{AB}(1 + p_M\delta r)p_b(P_{Mb} - P_{MB})$$
$$- r_{MA}p_m(P_{mB} - P_{MB}) + \frac{P_{MB}^2}{2}, \qquad (20)$$

where s_- and s_+ are given by Equation 15, while p_B and p_b are given by Equation 13. For example, the expected number of copies produced by an allele A present on background MB at t is $1 + s_A + s_B p_b$ (to the first order in selection coefficients). Of these, a proportion $r_{AB}(1 + \delta r \times p_M)p_b$ will move to the Mb background—since the expected recombination rate between selected loci, experienced by a chromosome carrying the M allele, is $r_{AB}(1 + \delta r \times p_M)$—a proportion $r_{MA}p_m$ will move to the mB background, and the rest will stay on the MB background (note that because we need to express the σ-coefficients of Equation 8 only to first order in selection coefficients and recombination rates, we can neglect double-recombination events). Allowing the modifier to affect the recombination rate r_{MA} would not affect the equations above. Specifically, we may denote the recombination rates between loci (M, m) and (A, a) in mm, Mm, and MM individuals by $r_{MA}(1 - \delta r_{MA})$, r_{MA}, and $r_{MA}(1 + \delta r_{MA})$, respectively. However, because recombination between (M, m) and (A, a) would affect genotype frequencies only when it occurs in Mm heterozygotes, δr_{MA} would not enter into the equations.

The system (20), with boundary conditions (19), can be solved numerically to obtain the P_j's for arbitrary t. In the case where A is present as a single copy at t, and assuming linkage equilibrium between the (M, m) and (B, b) loci at t, the expression for the Fourier transform ψ becomes (from Equation 18)

$$\psi = p_B(p_M e^{-P_{MB}} + p_m e^{-P_{mB}}) + p_b(p_M e^{-P_{Mb}} + p_m e^{-P_{mb}}). \qquad (21)$$

Using Equations 19–21, one can express ψ for any time t of occurrence of the A mutation.

In the following, we use scaled parameters and variables; this allows us to factor out selection coefficients and obtain more general results, in terms of the ratio of selection coefficients at both loci. In the diffusion limit,

ω_j and P_j are of order s_A for all backgrounds j. We define the scaled variables $\tilde{\omega}_j$ and $\tilde{P}_j$ as

$$\tilde{\omega}_j = \frac{\omega_j}{2s_A}, \quad \tilde{P}_j = \frac{P_j}{2s_A}. \qquad (22)$$

We then use the scaled parameters

$$T = s_B t, \quad T^* = s_B t^*, \quad \theta = \frac{s_A}{s_B}, \quad \rho_{MA} = \frac{r_{MA}}{s_B}, \quad \rho_{AB} = \frac{r_{AB}}{s_B}. \qquad (23)$$

The frequency of allele B as a function of time thus becomes $p_B = 1/(1 + e^{-T})$. Using these scaled parameters and variables, Equations 20 become

$$\frac{\partial \tilde{P}_{mb}}{\partial T} = -(\theta - p_B)\tilde{P}_{mb} - \rho_{AB}(1 - p_m\delta r)p_B(\tilde{P}_{mB} - \tilde{P}_{mb})$$
$$- \rho_{MA}p_M(\tilde{P}_{Mb} - \tilde{P}_{mb}) + 2\theta\tilde{P}_{mb}^2$$

$$\frac{\partial \tilde{P}_{Mb}}{\partial T} = -(\theta - p_B)\tilde{P}_{Mb} - \rho_{AB}(1 + p_M\delta r)p_B(\tilde{P}_{MB} - \tilde{P}_{Mb})$$
$$- \rho_{MA}p_m(\tilde{P}_{mb} - \tilde{P}_{Mb}) + 2\theta\tilde{P}_{Mb}^2$$

$$\frac{\partial \tilde{P}_{mB}}{\partial T} = -(\theta + p_b)\tilde{P}_{mB} - \rho_{AB}(1 - p_m\delta r)p_b(\tilde{P}_{mb} - \tilde{P}_{mB})$$
$$- \rho_{MA}p_M(\tilde{P}_{MB} - \tilde{P}_{mB}) + 2\theta\tilde{P}_{mB}^2$$

$$\frac{\partial \tilde{P}_{MB}}{\partial T} = -(\theta + p_b)\tilde{P}_{MB} - \rho_{AB}(1 + p_M\delta r)p_b(\tilde{P}_{Mb} - \tilde{P}_{MB})$$
$$- \rho_{MA}p_m(\tilde{P}_{mB} - \tilde{P}_{MB}) + 2\theta\tilde{P}_{MB}^2. \qquad (24)$$

From Equation 2, the change in frequency of the modifier after T^* can be obtained from $E[S^{\rho-1}D]$, where $\rho = r_{MA}/s_A = \rho_{MA}/\theta$, S is the total number of A alleles at T^*, and D is the linkage disequilibrium between loci (M, m) and (A, a) at T^*, multiplied by population size N (see APPENDIX A). The variables S and D are given by

$$S = n_{MB} + n_{mB}$$
$$D = p_m n_{MB} - p_M n_{mB}. \qquad (25)$$

In APPENDIX C, we present a method to derive $E[S^{\rho-1}D]$ from the Fourier transform ψ, in the case of a modifier of small effect (δr small), and when linkage between the modifier and the (A, a) locus is sufficiently tight ($\rho < 1$). We then calculate the selection gradient on the modifier allele M, defined as

$$s_M = \frac{1}{p_M p_m}\frac{\partial E[\Delta p_M]}{\partial \delta r}. \qquad (26)$$

The expected change in frequency of the modifier over the whole process is given by

$$E[\Delta p_M] = \delta r s_M p_M p_m + o(\delta r). \qquad (27)$$

Interestingly, APPENDIX C shows that s_M is independent of the initial frequency of the modifier (p_M) and can be written in the form $s_M = 2s_A\tilde{s}_M$, where $\tilde{s}_M$ is a function of T, θ, ρ_{AB}, ρ_{MA}, and Ns_A. A measure of the strength of
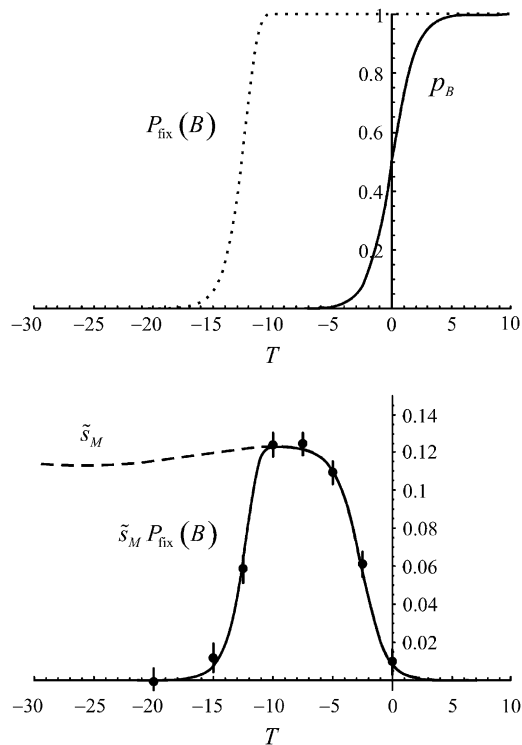
FIGURE 3.—(Top) Deterministic increase in frequency of allele B (solid curve) and probability of fixation of B (dashed curve, from Equation 28) from initial frequency p_B (given by Equation 13) as a function of time T. (Bottom) Scaled selection gradient on the modifier $\tilde{s}_M = s_M/(2s_A)$, as a function of T. The dashed curve shows the results obtained from Equation C23, and the solid curve shows the results obtained from (C23) multiplied by the probability of fixation of B (from Equation 28). Dots show simulation results (average over 10^7 replications); error bars measure ±1.96 SE. Parameter values are: $\theta = 0.1$, $\rho_{MA} = 0.01$, $\rho_{AB} = 0.2$, and $Ns_A = 10^4$; in the simulations $s_B = 0.1$, $\delta r = 0.5$, with $N = 10^6$ and $p_M = 0.5$, while in the analytical model the splice point T^* is set to 50.

selection at the modifier locus, generated by the Hill–Robertson effect at the selected loci, is thus given by $2s_A\delta r\tilde{s}_M$. As shown in APPENDIX C, obtaining $\tilde{s}_M$ for a given set of parameter values involves solving numerically a system of differential equations (Equations C15–C22) and performing a numerical integration (Equation C23). These operations have been implemented in a Mathematica notebook, which is available on request. Modifying our model to represent different orderings of the three loci is straightforward and leads to minor changes to our equations (see end of APPENDIX C).

Figure 3 shows a test of our method against simulations, for $\theta = 0.1$, $\rho_{MA} = 0.01$, $\rho_{AB} = 0.2$, and $Ns_A = 10^4$ (in the simulations $s_B = 0.1$, $\delta r = 0.5$, $N = 10^6$, and the initial frequency of the modifier is $p_M = 0.5$). In the simulation program, we start from a population where allele B is at frequency $p_B = 1/(1 + e^{-T})$ (the deterministic frequency of B at time T), allele M is at frequency p_M, and B and M are in linkage equilibrium.

Allele A is introduced as a single copy in a random background; then for each generation we calculate the expected change in genotype frequencies and sample the new frequencies according to a multinomial distribution. We stop when allele A has disappeared or has reached fixation and measure the change in frequency of M over the whole process. Figure 3 shows that our method provides a good prediction of selection on the modifier, as long as T is not too negative (bottom, dashed curve). The discrepancy for $T < \sim{-}10$ comes from the fact that the model assumes that allele B is established at T. However, when T is too negative, the initial frequency of B is too low for the allele to be established; this is shown in Figure 3, top, which plots the deterministic trajectory of allele B (solid curve) and its probability of fixation as a function of T (dashed curve), where we calculate the fixation probability from the standard one-locus diffusion formula

$$P_{\text{fix}}(B) = \frac{1 - e^{-2Ns_B p_B}}{1 - e^{-2Ns_B}} = \frac{1 - e^{-2\sigma_A p_B/\theta}}{1 - e^{-2\sigma_A/\theta}}, \qquad (28)$$

where $\sigma_A = Ns_A$ (e.g., CROW and KIMURA 1970). When $T < \sim{-}15$, Figure 3 shows that allele B has a high chance of being lost from the population, in which case selection on the modifier disappears. If we multiply the expression for the change in frequency of the modifier given by Equation C23 with the probability of fixation of B (given by Equation 28), we obtain a result that matches the simulations very well (Figure 3, bottom, solid curve). To generate the curves of Figure 3, we set the value of time T^* to 50; however, other values of T^* between 30 and 350 lead to indistinguishable curves. When $T^* < 30$, allele A can be still present in the population but not established (which leads to an overestimation of the change in frequency of the modifier), while when $T^* > 350$, allele A, when still present in the population, has already reached a substantial frequency (which may lead either to an underestimation or to an overestimation of the change in frequency of the modifier, depending on the parameter values, as the model ignores changes in modifier frequency before T^*, but tends to overestimate the change after T^*). The fact that the results are insensitive to T^* over a large range of values shows that the method is consistent. In Figure 4, we show that our model also gives very good results for different values of the selection coefficients s_A and s_B. Finally, Figure 5 compares results obtained using our method and Otto and Barton's method for $s_A = s_B = 0.1$, $r_{MA} = 0.001$ (Figure 4A), and $r_{MA} = 0.01$ (Figure 4B), corresponding to the values used in Figures 7 and 9 of OTTO and BARTON (1997). Figure 5 shows that the present method provides a better prediction of the change in frequency of the modifier. In Figure 5A, it can be seen that both the model and simulations give a bimodal curve; this effect is explained in the next section.
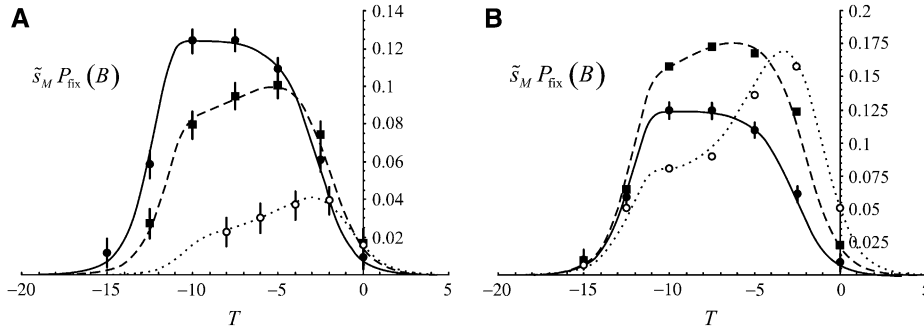
FIGURE 4.—Scaled selection gradient on the modifier $\tilde{s}_M = s_M/(2s_A)$ times the probability of fixation of B as a function of T. Curves: results obtained from Equation C23, multiplied by the probability of fixation of B (from Equation 28). Circles and squares: simulation results (average over 10^7 replications); error bars measure ± 1.96 SE (when the error bars do not appear, they are smaller than the dots). (A) $s_A = 0.01$ and $s_B = 0.1$ (solid curve, solid circles), 0.05 (dashed curve, squares), 0.02 (dotted curve, open circles). (B) $s_B = 0.1$ and $s_A = 0.01$ (solid curve, solid circles), 0.02 (dashed curve, squares), 0.05 (dotted curve, open circles). In each case $r_{MA} = 0.001$, $r_{AB} = 0.02$, $N = 10^6$. In the simulations $\delta r = 0.5$ and $p_M = 0.5$, while in the analytical model the splice point T^* is set to 50.

NET SELECTION GRADIENT OVER RECURRENT SWEEPS

What will be the net selection gradient for a modifier increasing the total map length of a genome, under a constant input of beneficial mutations? We assume here that the modifier increases the recombination rate uniformly over the genome and that adaptive sweeps are sufficiently rare that we can neglect interactions among more than two beneficial alleles. More specifically, we derive an expression for the net selection gradient to the order Λ^2, where Λ is the rate of adaptive substitutions, and neglect terms of higher order in Λ. Our model allows us to consider two classes of mutations. One can assume, for example, that some strongly selected alleles enter the population at a given rate (sufficiently low that the effect of interactions among these mutations can be neglected), while weakly selected mutations occur at a higher rate and have a reduced fixation probability due to the recurrent sweeps of strongly selected alleles. The interaction between strongly and weakly selected alleles will select for higher recombination rates. Interactions among weakly selected alleles will also select for more recombination, but as the effect of these interactions is weaker (*e.g.*, BARTON 1995b), they may be neglected (this should be checked, however, as this depends on how much more numerous weakly selected alleles are). To obtain the average effect of strong substitutions on weak ones, we can consider the case where allele B is

strongly selected and allele A weakly selected. Equation C23 then gives the rate of increase of the modifier, given that B fixes, and for given positions of the (A, a) and (B, b) loci. The selection gradient, conditional on fixing both alleles, is then given by $s_M/P_{\text{fix}}(A)$, where $P_{\text{fix}}(A)$ is the fixation probability of allele A, which can be obtained from Equation 6 in BARTON (1995b). Because $P_{\text{fix}}(A)$ can be written as $2s_A\tilde{P}_{\text{fix}}(A)$, where $\tilde{P}_{\text{fix}}(A)$ is a function of T, θ, and ρ_{AB}, while $s_M = 2s_A\tilde{s}_M$, the selection gradient given the fixation of both alleles is $\tilde{s}_M/\tilde{P}_{\text{fix}}(A)$, which depends on T, θ, ρ_{MA}, ρ_{AB}, and Ns_A. Averaging over time gives the net selection gradient, for given positions of alleles A and B:

$$\overline{s_M}(\rho_{MA}, \rho_{AB}) = \int_{-\infty}^{+\infty} \frac{\tilde{s}_M}{\tilde{P}_{\text{fix}}(A)} dt = \frac{1}{s_B} \int_{-\infty}^{+\infty} \frac{\tilde{s}_M}{\tilde{P}_{\text{fix}}(A)} dT. \quad (29)$$

The net selection gradient, under constant rates of substitutions Λ_A and Λ_B for weakly and strongly selected alleles, and for all possible positions of selected loci, is then obtained by integrating over the genetic map,

$$\overline{s_M} = \frac{\Lambda_A\Lambda_B}{R^2} \iint_0^{+\infty} \overline{s_M}(\rho_{MA}, \rho_{AB}) dr_{MA}\, dr_{AB}$$

$$= (s_B)^2 \frac{\Lambda_A\Lambda_B}{R^2} \iint_0^{+\infty} \overline{s_M}(\rho_{MA}, \rho_{AB}) d\rho_{MA}\, d\rho_{AB}, \quad (30)$$
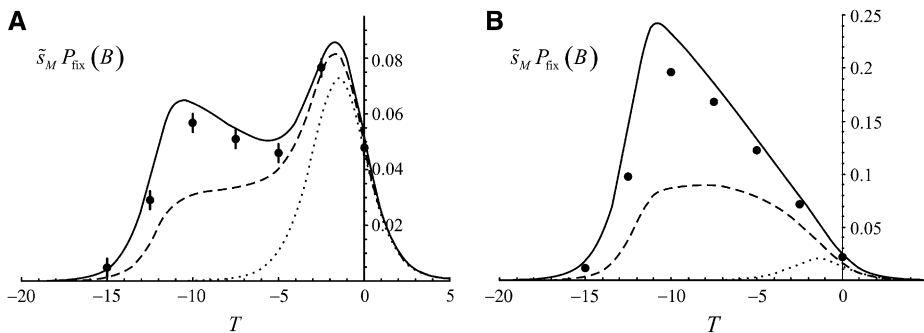

FIGURE 5.—Scaled selection gradient on the modifier $\tilde{s}_M = s_M/(2s_A)$ times the probability of fixation of B as a function of T. Solid curves, results obtained from Equation C23; dotted curves, results obtained from Equations 7–10 and 19 in OTTO and BARTON (1997); dashed lines, results obtained by the method of OTTO and BARTON (1997), taking into account the initial linkage disequilibrium between selected loci. Dots: simulations results (average over 2×10^7 replications), error bars measure ± 1.96 SE (in B, the error bars are smaller than the size of dots). Parameter values: $s_A = 0.1$, $s_B = 0.1$, $r_{AB} = 0.02$, $N = 10^6$, and $r_{MA} = 0.001$ (A) and 0.01 (B). In the simulations, $\delta r = 0.5$ and $p_M = 0.5$.

where R is the total map length (because only selected loci that are sufficiently tightly linked to the modifier contribute significantly to its increase in frequency, we can take $+\infty$ as the integration limit when integrating over the genetic map). Equations 29 and 30 finally give

$$\overline{s_M} = s_B \frac{\Lambda_A \Lambda_B}{R^2} \iint_0^{+\infty} \int_{-\infty}^{+\infty} \frac{\tilde{s}_M}{\tilde{P}_{\text{fix}}(A)} dT \, d\rho_{MA} \, d\rho_{AB}, \quad (31)$$

where the triple integral depends on θ and on Ns_A. Moreover, we expect that $\overline{s_M}$ should scale roughly with $1/\log(Ns_A)$. Indeed, from Equation C23, $\tilde{s}_M$ takes the form $(2Ns_A)^{-\rho} = e^{-\rho \log(2Ns_A)}$, multiplied by a function of ρ_{MA}, ρ_{AB}, and θ, say $f(\rho_{MA}, \rho_{AB}, \theta)$. Expanding f using a Taylor series and integrating over $\rho = \rho_{MA}/\theta$ gives

$$\int_0^\infty e^{-\rho \log(2Ns_A)} f(\rho_{MA}, \rho_{AB}, \theta) d\rho$$
$$= \int_0^\infty e^{-\rho \log(2Ns_A)} [f(0, \rho_{AB}, \theta) + \rho f'(0, \rho_{AB}, \theta) + \cdots] d\rho$$
$$= \frac{f(0, \rho_{AB}, \theta)}{\log(2Ns_A)} + \frac{f'(0, \rho_{AB}, \theta)}{[\log(2Ns_A)]^2} + \cdots, \quad (32)$$

which is dominated by the first term when $\log(2Ns_A)$ is large. This slow decrease in the strength of selection on the modifier as population size increases probably reflects the fact that what matters most in a large population are stochastic effects occurring when selected alleles are present in small numbers of copies, which are not affected much by population size. Finally, it is important to note that the calculation of $\overline{s_M}$ given above considers only pairwise interactions among selected loci and thus cannot be applied to the case where more than two selective sweeps overlap in time. However, as is discussed later, rates of amino acid substitutions in the human and Drosophila lineages are rather low, suggesting that interactions between more than two sweeps may be quite rare.

Equation 31 involves the selection gradient for the modifier, conditional upon the fixation of allele B ($\tilde{s}_M$). Moreover, the integral over T includes cases where allele A enters the population before allele B. In the previous section, we compared simulation results with the model's predictions for the unconditional selection gradient for the modifier, $\tilde{s}_M P_{\text{fix}}(B)$. If we now condition upon the fixation of B, selection on the modifier can be substantial even when allele A occurs before allele B, but is still segregating at the time where B occurs. This is shown, for example, in Figure 6, which plots the selection gradient for the modifier, conditional upon the fixation of allele B ($\tilde{s}_M$). The solid and dashed curves give the model predictions for $\rho_{MA} = 0.01$ and $\rho_{MA} = 0.005$ (respectively), obtained from Equation C23; dots represent simulation results, for $s_B = 0.1$. In the simulation program, allele B appears as a single copy at the time where $p_B = 1/N$ according to the deterministic expression (13); this time is ~ -13.8 for the parameter values used in Figure 6. For values of $T < -13.8$, allele A enters the population before
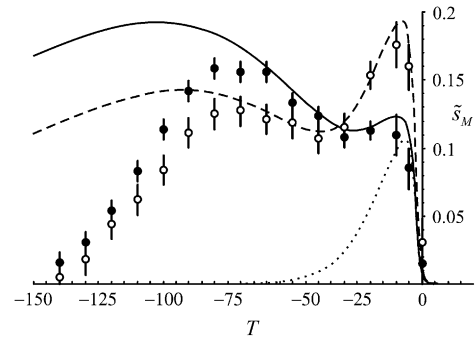

FIGURE 6.—Scaled selection gradient for the modifier ($\tilde{s}_M = s_M/(2s_A)$) as a function of the time of appearance of allele A (T), conditional on the fixation of allele B. Curves: results obtained from Equation C23, with $\rho_{MA} = 0.01$ (solid line) and $\rho_{MA} = 0.005$ (dashed line). Dotted curve: result obtained from Equations 7–10 and 19 in OTTO and BARTON (1997), with $\rho_{MA} = 0.01$. Dots: simulation results (average over 6×10^7 replications), with $\rho_{MA} = 0.01$ (solid circles) and $\rho_{MA} = 0.005$ (open circles); error bars measure ± 1.96 SE. In the simulations, allele B enters the population as a single copy (on a random background) at time $T = -13.8$. Other parameters: $\theta = 0.1$, $\rho_{AB} = 0.2$, $Ns_A = 10^4$; in the simulations, $s_B = 0.1$, $\delta r = 0.5$, $N = 10^6$, and $p_M = 0.5$. In the analytical model the splice point T^* is set to 50.

allele B and appears on a random background at the modifier locus. Then, at time -13.8, allele B enters the population on a random background. The program stops when both selected loci are fixed, and the change in frequency of the modifier is measured among those cases where B has reached fixation. For values of $T > -13.8$, allele B enters the population before allele A; again, the program runs until both loci are fixed, and the change in frequency of the modifier is measured only among cases where B is fixed.

Several aspects of Figure 6 deserve attention. First, the model does not provide a good prediction of the change in frequency of the modifier for very negative values of T, that is, when allele A appears long before allele B. This is due to the fact that the model assumes that, at a time T^* after the sweep of allele B, allele A is either lost or established, but still at low frequency in the population. However, for very negative values of T, A may have reached a high frequency at time T^* or even be fixed. For example, for $T = -125$, simulations indicate that the frequency of A at time $T^* = 50$, given that A will fix, is ~ 0.99. In that case, the branching process approximation (independent replication of the different copies of A) has ceased to apply long before T^* and probably even before the sweep of B. In this case, most of the change at the modifier locus occurs before T^*, while the model assumes no change before T^*. However, the model greatly overestimates the change after T^*, because it overestimates $E[S^{\rho-1}D]$ at T^* (this is confirmed by simulations), and also because Equation A7 assumes that allele A is still in low frequency at T^*. Another potential problem is that the model allows $p_B > 0$ before

allele B actually appears in the population (that is, before time -13.8). However, this does not seem to cause problems when A appears at time -50, for example. In any case, it appears unlikely that these problems can be solved by simple adjustments of our model: indeed, the discrepancy occurs when the sweeps of A and B overlap, which requires using the full three-locus recursions for describing the deterministic phase. We are currently working on an alternative method to treat this case, as is discussed next.

A second important result shown by Figure 6 is that the curves giving the selection gradient on the modifier as a function of T are bimodal (this can also be seen in Figure 5A and is confirmed by the simulations). This is due to the fact that the modifier has two different effects on the spread of allele A, whose relative importance depends on the time of appearance of A relative to B. When allele A appears after or shortly before allele B, its probability of fixation is reduced by the sweep of B. By reducing this interference, the recombination modifier increases the fixation probability of A and then benefits from the sweep of A (by hitchhiking). However, for values of $T < \sim -40$, the sweep of allele B has very little effect on the probability of fixation of A (this can be seen, for example, in Figure 4 of OTTO and BARTON 1997): A is either lost or already established when B enters the population. In this case, the modifier does not affect the probability of fixation of A, but still has an effect in accelerating the sweep of A, by helping it to recombine onto the B background. This effect also generates an indirect selective force on the modifier and explains the leftmost maxima of the curves of Figure 6. Because the two peaks are separated by a time $\sim(1/s_A)\log(N)$ (which is the time needed for allele A to become established, in the absence of allele B), they will move apart as N increases (this is confirmed by simulations—not shown). It is possible in principle to calculate the advantage of the modifier via accelerating sweeps (left peak) using a deterministic calculation; indeed, one can relatively easily describe changes in genotype frequencies when A and B are rare and then solve for the change when they become common using eight coupled differential equations, describing evolution at the three loci. Such a calculation is presently being assembled (N. H. BARTON and D. ROZE, unpublished data).

Finally, Figure 6 shows that increasing the recombination rate ρ_{MA} can have different qualitative effects on the rate of increase of the modifier, depending on the time of appearance of A relative to B (this can also be seen by comparing Figure 5A and 5B). When A appears after or shortly before B, increasing ρ_{MA} decreases the rate of increase of the modifier, by decreasing the effect of hitchhiking. When A appears long before B, however, increasing ρ_{MA} has the opposite effect. Before the time when B enters the population, the only effect of recombination between loci (M, m) and (A, a) is to

decrease the variance in linkage disequilibrium D_{MA}. By using the method of BARTON and OTTO (2005), we found that the only effect of an initial variance in D_{MA} is to reduce the genetic variance at the modifier locus, $p_M p_m$. Indeed, in the absence of selection at locus (B, b), the expected change in genetic variance at the modifier locus, over one generation, is given by

$$E[\, p'_M p'_m] = E[p_M p_m] - 4(s_A)^2 E[(D_{MA})^2] + o[(s_A)^2].$$

(33)

This corresponds to the classical result that selection tends to reduce genetic diversity at linked loci. Because the expected change in frequency of the modifier, at any generation, is proportional to the genetic variance at the modifier locus (at this generation), the variance in D_{MA} tends to reduce the rate of increase of the modifier. Increasing the recombination rate ρ_{MA} decreases the variance in D_{MA} and thus decreases this effect.

Because the present method does not deal well with the case where allele A occurs a long time before allele B (as shown by Figure 6), it has to be coupled to another method describing this case better before we can use Equation 31 to obtain the net selection gradient for the modifier under recurrent sweeps (N. H. BARTON and D. ROZE, unpublished data). Nevertheless, we can still obtain some results for the case where $\theta = 1$ ($s_A = s_B$). Indeed, simulations indicate that the method gives accurate predictions as long as A occurs after B (e.g., Figure 5). However, in the opposite case (when A occurs before B), and when $s_A = s_B$, we can simply swap A and B and continue to use the same method. However, to do so we need to incorporate the fact that allele B starts from frequency $1/N$; we thus need to specify the values of N and s_A, while Equation 31 depends only on the product Ns_A. We performed numerically the triple integration of Equation 31 for the case where $s_A = s_B = 0.05$ and obtained $s_M R^2/\Lambda^2 \approx 0.151, 0.126$, and 0.102 for $N = 10^5$, $N = 10^6$, and $N = 10^7$, respectively. These results confirm the fact that the selection gradient roughly scales with $1/\log(Ns_A)$ for large $\log(Ns_A)$; indeed, $s_M \times \log(Ns_A) \cdot R^2/\Lambda^2$ gives $1.29, 1.36$, and 1.34 for $N = 10^5, 10^6$, and 10^7, respectively. The values given above lead to exceedingly small selection gradients, using current estimates for Λ and R for Drosophila and humans. From comparisons between *D. simulans* and *D. yakuba*, the rate of adaptive amino acid substitutions has been estimated at one every 450 generations (SMITH and EYRE-WALKER 2002). This estimate has since been decreased to one every 800 generations (BIERNE and EYRE-WALKER 2004). The total map length R is estimated to be $\sim$3.6 for *D. simulans* (TRUE *et al.* 1996). For humans, the rate of adaptive amino acid substitution is estimated to be one every 200 years since the divergence with old-world monkeys (FAY *et al.* 2001), which, using a generation time of 25 years (EYRE-WALKER and KEIGHTLEY 1999) gives $\Lambda = \frac{1}{8}$ per generation. The total map length is $\sim R = 34$ for

humans. In both cases, the ratio Λ^2/R^2 is very small (of order 10^{-5} for humans and 10^{-7} for Drosophila). Combining the present method with the method of N. H. BARTON and D. ROZE (unpublished data) that studies the effect of recombination modifiers in accelerating sweeps and the method of BARTON and OTTO (2005) that deals with the effects of small departures from deterministic trajectories once alleles are common, we should be able to obtain estimates over a range of selection coefficients and population sizes. However, it appears unlikely that the selection gradient for recombination modifiers will be large, except during periods of frequent selective sweeps or in genomes with very tight linkage.

DISCUSSION

Apart from the idea that recombination brings a mechanistic advantage to individuals in terms of DNA repair (*e.g.*, BERNSTEIN *et al.* 1988), the different hypotheses proposed to explain the maintenance of high rates of recombination in higher organisms state that recombination is advantageous because it reduces LD among selected loci (KONDRASHOV 1993; BARTON and CHARLESWORTH 1998; OTTO and LENORMAND 2002). Different possible sources of LD have been identified. First, LD may be generated by epistatic interactions among loci. In this case, increased recombination is selected when epistasis is weakly negative (FELDMAN *et al.* 1980; KONDRASHOV 1982, 1988; BARTON 1995a) and not too variable among pairs of loci (OTTO and FELDMAN 1997) or when epistasis fluctuates over short periods of time (CHARLESWORTH 1976; BARTON 1995a). However, experimental evidence does not indicate any strong trend for epistasis to follow this pattern, and it seems likely that the variance in epistasis is high in most cases (*e.g.*, SEAGER and AYALA 1982; SEAGER *et al.* 1982; DE VISSER *et al.* 1996, 1997a,b; ELENA and LENSKI 1997; ELENA 1999; DE LA PEÑA *et al.* 2000; WLOCH *et al.* 2001). A second possible source of LD is environmental heterogeneity, in which case recombination is selected when directional selection at different loci covaries negatively between habitats (LENORMAND and OTTO 2000). Finally, LD may be generated by drift and directional selection: in a finite population, some negative disequilibrium does develop, on average, between selected loci, even in the absence of epistasis (HILL and ROBERTSON 1966; FELSENSTEIN 1974). By decreasing the effect of these interactions among loci, recombination increases the rate of fixation of advantageous alleles (BARTON 1995b). In this article, we presented a new method to quantify the strength of selection acting on a recombination modifier, due to this last effect.

Although the idea that sex and recombination increase the rate of adaptation has a long history in population genetics, theoretical models addressing the question of the effect of interference among loci on the evolution of recombination have been developed relatively late. OTTO and BARTON (1997) calculated the expected change in frequency of a modifier allele affecting the recombination rate between two selected loci in a finite population, using a method derived from branching processes. This model has led to important insights, but it does not always provide good quantitative estimates of the strength of selection at the modifier locus. This is due to the fact that interference among selected loci is neglected in part of the analysis: the linkage disequilibrium generated by the Hill–Robertson effect is taken into account in the calculation of probabilities of fixation of selected alleles, but not in the calculation of the hitch given by a selected allele to the recombination modifier, given that it fixes. In a more recent work, BARTON and OTTO (2005) used a different method to calculate the effect of linkage disequilibrium generated by drift on the change in frequency of a recombination modifier; however, their analysis assumes small departures from deterministic trajectories and thus does not apply when selected alleles start from a small number of copies (which is the case where drift will have the strongest effect, when population size is large).

In this article, we have presented a new method that overcomes these difficulties, in the case where population size is sufficiently large. The essence of the method lies in the calculation of the probability distribution of the number of copies of a new selected allele on different genetic backgrounds, once it is established but while it is still rare. Assuming a large population size, one can then neglect further stochastic fluctuations and express the change in frequency of the recombination modifier in terms of this probability distribution. Because the new allele remains rare during the establishment phase, the change in number of its copies on the different backgrounds can be described by a branching process, which generates a system of differential equations that can be solved numerically to obtain the Fourier transform of the probability distribution. This method fully takes into account the linkage disequilibrium between selected loci generated by the initial mutation event, by genetic drift during the establishment phase, and by the fact that the allele tends to be found more often in coupling with the other selected allele on the paths leading to its fixation than on the paths leading to its loss. We used this method to calculate the full probability distribution of the number of copies of a new selected allele when selection also occurs at a linked locus and found a good match with simulation results (Figure 1). To obtain an expression for the change in frequency of the recombination modifier, we coupled this stochastic approach to a deterministic calculation of the hitch given to the modifier by the new selected allele, after establishment. By comparing our results with simulations, we found

that this method provides very accurate predictions as long as the basic assumptions of the model are satisfied, that is, when allele *A* appears once allele *B* is already established.

As illustrated by Figure 6, our method also provides good results when the weakly selected allele *A* enters the population before the strongly selected allele *B* (both alleles being initially present as single copies), as long as *A* occurs not too long before *B*. Furthermore, Figure 6 shows that the modifier increases in frequency through two different mechanisms, whose relative importance depends on the relative times of occurrence of the two favored alleles: the modifier increases the probability of fixation of allele *A* (this effect being marked mostly when *A* occurs around the time of appearance of *B*), and it also accelerates the sweep of *A* (this effect being marked mostly when allele *A* is already established when allele *B* occurs). These two effects generate an indirect pressure selecting for the modifier. We have also seen that in some cases (when allele *A* occurs before allele *B*) increasing the recombination rate r_{MA} can lead to a greater expected increase in frequency of the modifier. This is due to the fact that increasing r_{MA} decreases the variance in linkage disequilibrium D_{MA}. Indeed, the variance in D_{MA} tends to reduce the genetic variance at the modifier locus, which in turn reduces the expected change in frequency of the modifier.

Finally, we have seen that the net selection gradient for a modifier affecting recombination over a whole genome, during recurrent selective sweeps, takes the form $s_B \Lambda_A \Lambda_B / R^2$ (where Λ_A and Λ_B are rates of substitutions and R is the map length of the genome), multiplied by a function of $\theta = s_A / s_B$ and of Ns_A; this function has to be obtained by integrating numerically over time and over the possible locations of selected loci. Because our method does not describe accurately the case where allele *A* occurs a long time before allele *B*, however, we could perform this integration only for the case where $s_A = s_B$. For $s_A = s_B = 0.05$, we obtained that the net selection gradient for the modifier is given by $s_M R^2 / \Lambda^2 \approx 0.151, 0.126$, and 0.102 for $N = 10^5, 10^6$, and 10^7, respectively. Combining the present method with the method of N. H. BARTON and D. ROZE (unpublished data), however, we should be able to obtain more results for different values of θ and of Ns_A.

According to current estimates of rates of amino acid substitutions for Drosophila and humans, the opportunity for indirect selection on recombination rates is rather small. A much higher estimate for the rate of adaptive substitution (of about one every 10 generations) has been advanced for Drosophila from a study of differences in genetic variability across loci (SCHLÖTTERER *et al.* 1997; NURMINSKY 2001). While other factors than selective sweeps may also account for these differences, it remains possible that a proportion of advantageous mutations do not cause any change in protein composition (mutations in regulatory sequences, for ex-

ample). In a multilocus simulation study, OTTO and BARTON (2001) obtained substantial changes in recombination rates due to Hill–Robertson effects between selected loci (in particular, Figure 4 in OTTO and BARTON 2001); similar results were obtained in much larger populations that are spatially structured (MARTIN *et al.* 2006). These results are not incompatible with the present results, however, as sweeps were occurring at several loci at the same time (high Λ), in a small-sized genome (small R), while population size was relatively small ($N \leq 1000$). Such conditions (high Λ, small N) may explain increases in recombination rates obtained after selection experiments. Similarly, ILES *et al.* (2003) found substantial increases in recombination when selective sweeps occur simultaneously at more than two loci. However, such a situation may appear unlikely given the current estimates of rates of amino acid substitutions. Still, this does not mean that Hill–Robertson interactions have little effect on the evolution of sex and recombination. First, it is possible that adaptation occurs mostly during relatively short time periods, during which substitution rates are higher. Also, the direction of selection may fluctuate over time, in which case rates of substitutions may not be a good indicator of the frequency of sweeps occurring within populations. Second, population structure increases the effects of Hill–Robertson interference among loci (MARTIN *et al.* 2006). Third, the effect of deleterious alleles at mutation–selection balance may be far greater than the effect of beneficial alleles sweeping through populations: indeed, given the high rates of deleterious mutation that have been estimated for several species, populations should be polymorphic for deleterious mutations at many loci. The Hill–Robertson effect will also generate negative linkage disequilibria among these loci, reducing the efficiency of selection against deleterious alleles. More analytical and simulation work is thus needed before we can assess the plausibility of the hypothesis that stochastic interactions among loci play an important role in the evolution of sex and recombination.

LITERATURE CITED

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.

BARTON, N. H., 1998 The effect of hitch-hiking on neutral genealogies. Genet. Res. **72:** 123–133.

BARTON, N. H., and B. CHARLESWORTH, 1998 Why sex and recombination? Science **281:** 1986–1990.

BARTON, N. H., and S. P. OTTO, 2005 Evolution of recombination due to random drift. Genetics **169:** 2353–2370.

Barton, N. H., and M. Turelli, 1991 Natural and sexual selection on many loci. Genetics **127:** 229–255.

Bernstein, H., F. A. Hopf and R. E. Michod, 1988 Is meiotic recombination an adaptation for repairing DNA, producing genetic variation, or both?, pp. 139–160 in *The Evolution of Sex: An Examination of Current Ideas*, edited by R. E. Michod and B. R. Levin. Sinauer Associates, Sunderland, MA.

Bierne, N., and A. Eyre-Walker, 2004 The genomic rate of adaptive amino-acid substitution in *Drosophila*. Mol. Biol. Evol. **21:** 1350–1360.

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

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

Charlesworth, B., and D. Charlesworth, 1985b Genetic variation in recombination in *Drosophila*. II. Genetic analysis of a high recombination stock. Heredity **54:** 85–98.

Crow, J. F., and M. Kimura, 1965 Evolution in sexual and asexual populations. Am. Nat. **99:** 439–450.

Crow, J. F., and M. Kimura, 1970 *An Introduction to Population Genetics Theory.* Harper & Row, New York.

de la Peña, M., S. F. Elena and A. Moya, 2000 Effect of deleterious mutation-accumulation on the fitness of RNA bacteriophage MS2. Evolution **54:** 686–691.

de Visser, J. A. G. M., R. F. Hoekstra and H. van den Ende, 1996 The effect of sex and deleterious mutations on fitness in *Chlamydomonas*. Proc. R. Soc. Lond. Ser. B **263:** 193–200.

de Visser, J. A. G. M., R. F. Hoekstra and H. van den Ende, 1997a An experimental test for synergistic epistasis and its application in Chlamydomonas. Genetics **145:** 815–819.

de Visser, J. A. G. M., R. F. Hoekstra and H. van den Ende, 1997b Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. Evolution **51:** 1499–1505.

Elena, S. F., 1999 Little evidence for synergism among deleterious mutations in a nonsegmented RNA virus. J. Mol. Evol. **49:** 703–707.

Elena, S. F., and R. E. Lenski, 1997 Test of synergistic interactions among deleterious mutations in bacteria. Nature **390:** 395–397.

Eyre-Walker, A., and P. D. Keightley, 1999 High genomic deleterious mutation rates in hominids. Nature **397:** 344–347.

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

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

Feller, W., 1951 Diffusion processes in genetics. Proc. Second Berkeley Symp. Math. Stat. Prob. 227–246.

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

Felsenstein, J., and S. Yokohama, 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.

Harris, T. E., 1963 *The Theory of Branching Processes.* Dover Phoenix Editions, New York.

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

Iles, M. M., K. Walters and C. Cannings, 2003 Recombination can evolve in large finite populations given selection on sufficient loci. Genetics **165:** 2249–2258.

Kidwell, M. G., 1972 Genetic change in recombination value in *Drosophila melanogaster*. I. Selection for high and low recombina-

tion and some properties of recombination-modifying genes. Genetics **70:** 419–432.

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

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

Kondrashov, A. S., 1993 Classification of hypotheses on the advantage of amphimixis. J. 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.

Korol, A. B., I. A. Preygel and S. I. Preygel, 1994 *Recombination Variability and Evolution. Algorithms of Estimation and Population-Genetic Models:* Chapman & Hall, London.

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

Lighthill, M. J., 1958 *An Introduction to Fourier Analysis and Generalised Functions.* Cambridge University Press, Cambridge, UK.

Martin, G., S. P. Otto and T. Lenormand, 2006 Selection for recombination in structured populations. Genetics **172:** 593–609.

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

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

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

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

Ohta, T., and M. Kimura, 1969 Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. Genetics **63:** 229–238.

Ohta, T., and M. Kimura, 1971 Linkage disequilibrium between two segregating nucleotide sites under the steady flux of mutations in a finite population. Genetics **68:** 570–580.

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

Otto, S. P., and N. H. Barton, 2001 Selection for recombination in small populations. Evolution **55:** 1921–1931.

Otto, S. P., and M. W. Feldman, 1997 Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theor. Popul. Biol. **51:** 134–147.

Otto, S. P., and T. Lenormand, 2002 Resolving the paradox of sex and recombination. Nat. Rev. Genet. **3:** 252–261.

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

Seager, R. D., and F. J. Ayala, 1982 Chromosome interactions in *Drosophila melanogaster*. I. Viability studies. Genetics **102:** 467–483.

Seager, R. D., F. J. Ayala and R. W. Marks, 1982 Chromosome interactions in *Drosophila melanogaster*. II. Total fitness. Genetics **102:** 485–502.

Smith, N. G., and A. Eyre-Walker, 2002 Adaptive protein evolution in Drosophila. Nature **415:** 1022–1024.

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.

Wloch, D. M., R. H. Borts and R. Korona, 2001 Epistatic interactions of spontaneous mutations in haploid strains of the yeast *Saccharomyces cerevisiae*. J. Evol. Biol. **14:** 310–316.

APPENDIX A

We express here the total change in frequency of the modifier after t^*, for given numbers of A alleles present on the M and m backgrounds at t^*. We assume that the change in genotype frequencies is deterministic after t^*; in this case, the change in p_M and p_A (the frequencies of alleles M and A) over one generation can be written in the form

$$\Delta p_M = a_A D_{MA}, \Delta p_A = a_A p_A q_A \tag{A1}$$

(*e.g.*, BARTON and TURELLI 1991), where D_{MA} is the linkage disequilibrium between the two loci, a_A is a function of s_A and p_A, and $q_A = 1 - p_A$. We thus have

$$\Delta p_M = \Delta p_A \frac{D_{MA}}{p_A q_A}. \tag{A2}$$

Furthermore, it can be shown that $D_{MA}/(p_A q_A)$ decays at a rate $1 - r_{MA}$ per generation, so that

$$\left(\frac{D_{MA}}{p_A q_A}\right)_t = (1 - r_{MA})^{t-t^*}\left(\frac{D_{MA}}{p_A q_A}\right)_{t^*} \approx e^{-r_{MA}(t-t^*)}\frac{D}{S}, \tag{A3}$$

where D is the linkage disequilibrium D_{MA} at time t^* multiplied by population size N, and S is the number of copies of allele A in the population at time t^*. Furthermore, the dynamics of selection give

$$\left(\frac{p_A}{q_A}\right)_t \approx e^{s_A(t-t^*)}\left(\frac{p_A}{q_A}\right)_{t^*} \approx e^{s_A(t-t^*)}\frac{S}{N}, \tag{A4}$$

giving

$$e^{-r_{MA}(t-t^*)} = \left(\frac{S}{N}\right)^{\rho}\left(\frac{p_A}{q_A}\right)_t^{-\rho} \tag{A5}$$

with $\rho = r_{MA}/s_A$. We thus have

$$\frac{D_{MA}}{p_A q_A} = N^{-\rho} S^{\rho-1} D\left(\frac{p_A}{q_A}\right)^{-\rho}. \tag{A6}$$

The total change in frequency of allele M is then obtained by integrating over p_A (going from 0 to 1):

$$\Delta p_M = N^{-\rho} S^{\rho-1} D \int_0^1 \left(\frac{p_A}{q_A}\right)^{-\rho} dp_A = N^{-\rho}\Gamma(1-\rho)\Gamma(1+\rho)S^{\rho-1}D \tag{A7}$$

(BARTON 1998). For small ρ, $\Gamma(1+\rho)\Gamma(1-\rho)$ is close to 1. The average change in frequency of the modifier over all possible realizations at t^* is thus obtained by averaging the quantity $S^{\rho-1}D$ over the probability distribution ϕ at t^*.

APPENDIX B

Inverting the Fourier transform: We explain here how to derive the probability distribution of the number of A alleles (n) at t^* from the Fourier transform $\psi(\omega)$, in the two-locus model. Because n can take only discrete values, the true probability distribution of n is discrete. We explained in the text how one can obtain the probability of loss of allele A ($n = 0$) from the value of ψ for large ω. Here we denote $\phi_{>0}^D(n)$ the true (and discrete) probability distribution of n at time t^*, for $n > 0$. Because we assume that allele A is still rare at time t^*, we can choose an integer n_{max} such that $\phi_{>0}^D(n)$ becomes vanishingly small for $n > n_{max}$ (this allows us to reduce computation time). In the following we assume that n_{max} is an even number.

The discrete Fourier transform (DFT) of $\phi_{>0}^D$ is given by

$$\psi_{>0}^D(k) = \sum_{n=1}^{n_{max}} \phi_{>0}^D(n)e^{-2\pi\iota kn/n_{max}}. \tag{B1}$$

The DFT $\psi_{>0}^D$ takes n_{max} distinct values (for $k = 0, \ldots, n_{max} - 1$) and is periodic: for any integer j, we have $\psi_{>0}^D(k + jn_{max}) = \psi_{>0}^D(k)$ (periodicity comes from the fact that $e^{\iota x} = \cos x + \iota \sin x$). The sum in Equation B1 can also be written

$$\sum_{n=1}^{n_{max}} \phi_{>0}^D(n)e^{-\iota\omega_k n}, \quad \text{with } \omega_k = 2\pi\frac{k}{n_{max}}, \tag{B2}$$

and can be approximated by the integral

$$\int_1^{n_{max}} \phi_{>0}(n)e^{-\iota\omega_k n} dn = \psi_{>0}(\omega_k). \tag{B3}$$

The function $\psi_{>0}$ is the continuous transform of the distribution $\phi(n)$ for $n > 0$ and is given by $\psi(\omega) - \psi(\infty)$, where $\psi(\infty)$ is the limit of $\psi(\omega)$ as ω tends to infinity (see text). The distribution $\phi_{>0}^D$ can be recovered from $\psi_{>0}^D$ by the inverse DFT:

$$\phi_{>0}^D(n) = \frac{1}{n_{\max}} \sum_{k=0}^{n_{\max}-1} \psi_{>0}^D(k) e^{2\pi \iota k n / n_{\max}}. \tag{B4}$$

Because $\psi_{>0}^D$ is periodic, this is also

$$\phi_{>0}^D(n) = \frac{1}{n_{\max}} \sum_{k=-n_{\max}/2}^{n_{\max}/2-1} \psi_{>0}^D(k) e^{2\pi \iota k n / n_{\max}} \approx \frac{1}{n_{\max}} \sum_{k=-n_{\max}/2}^{n_{\max}/2-1} \psi_{>0}(\omega_k) e^{\iota \omega_k n}. \tag{B5}$$

The number of terms to be evaluated can then be reduced by noting that the real parts of $\psi_{>0}(\omega_k)$ and $\psi_{>0}(-\omega_k)$ are equal, while their imaginary parts cancel out (this can be seen from Equation B3). Thus, to invert the transform it is sufficient to calculate the real part of $\psi_{>0}(\omega_k) e^{\iota \omega_k n}$ for $n_{\max}/2 + 1$ values of ω_k between 0 and π,

$$\phi_{>0}^D(n) \approx \frac{1}{n_{\max}} \mathrm{Re}\left[\psi_{>0}(0) + (-1)^n \psi_{>0}(\pi) + 2 \sum_{k=1}^{n_{\max}/2-1} \psi_{>0}(\omega_k) e^{\iota \omega_k n} \right], \tag{B6}$$

where $\mathrm{Re}[x]$ stands for the real part of x. Computations could in principle be made more efficient by using Fast Fourier algorithms, although we did not explore this possibility here.

Calculating moments of ϕ: Moments of the probability distribution $\phi(n)$ can be obtained directly from its Fourier transform ψ. For example, the expected number of copies of allele A at time t^* is given by

$$M \equiv E[n] = \iota \frac{d\psi}{d\omega}\bigg|_{\omega=0}. \tag{B7}$$

Using Equation 12, and noting that P_B and $P_b = 0$ when $\omega = 0$, we obtain

$$M = -\iota \left(p_B \frac{dP_B}{d\omega} + p_b \frac{dP_b}{d\omega} \right), \tag{B8}$$

the derivatives being evaluated at $\omega = 0$. Differential equations of the derivatives of P_B and P_b with respect to ω are obtained by differentiating Equations 14, giving

$$\partial_t \frac{dP_b}{d\omega} = -(s_-)\frac{dP_b}{d\omega} + r_{AB} p_B \left(\frac{dP_b}{d\omega} - \frac{dP_B}{d\omega} \right)$$
$$\partial_t \frac{dP_B}{d\omega} = -(s_+)\frac{dP_B}{d\omega} + r_{AB} p_b \left(\frac{dP_B}{d\omega} - \frac{dP_b}{d\omega} \right). \tag{B9}$$

From this, and noting that $\partial_t p_B = s_B p_B p_b$, one obtains

$$\partial_t M = -s_A M \tag{B10}$$

and thus $M = e^{s_A(t^*-t)}$, corresponding to the deterministic increase of allele A. Our method thus does not detect the effect of interference on the expected number of copies of A at t (because changes in frequencies are expressed only to first order); however, it shows an effect on higher moments of the distribution ϕ. For example, the variance of the number of copies of A at t^* is given by

$$V \equiv E[n^2] - E[n]^2 = -\frac{d^2\psi}{d\omega^2} + \left(\frac{d\psi}{d\omega} \right)^2, \tag{B11}$$

the derivatives being evaluated at $\omega = 0$. After differentiating Equations 14 twice with respect to ω and simplifying, one obtains

$$\partial_t V = p_B \left(\frac{dP_B}{d\omega} \right)^2 + p_b \left(\frac{dP_b}{d\omega} \right)^2 + (s_A + 2r_{AB}) p_B p_b \left(\frac{dP_B}{d\omega} - \frac{dP_b}{d\omega} \right)^2 + s_B p_B p_b \left[\left(\frac{dP_B}{d\omega} \right)^2 - \left(\frac{dP_b}{d\omega} \right)^2 \right] - s_A V. \tag{B12}$$

Equations B9 and B12 can be solved numerically to express V as a function of t. Figure 2 illustrates that selection at the (B, b) locus increases the variance of the number of copies of A at t^*.

APPENDIX C

Here we develop a method to obtain the expected change in frequency of the recombination modifier, in the three-locus model. Equation 2 expresses this change in frequency in terms of the expectation of $S^{\rho-1}D$ at time t^*, where S and D are given by

$$S = n_{MB} + n_{mB}$$
$$D = p_m n_{MB} - p_M n_{mB}. \tag{C1}$$

Changing variables in Equation 17 (by expressing n_{MB} and n_{mB} in terms of S and D), we can express the Fourier transform of the probability distribution $\phi(S, D)$ as

$$\psi(\omega_S, \omega_D) = \int\int \phi(S, D) e^{-\iota\omega_S S - \iota\omega_D D} dS\, dD \tag{C2}$$

with

$$\omega_S = p_M \omega_{MB} + p_m \omega_{mB}$$
$$\omega_D = \omega_{MB} - \omega_{mB}. \tag{C3}$$

Using these new variables, the boundary conditions (19) become

$$P_{MB} = \iota(\omega_S + p_m \omega_D)$$
$$P_{mB} = \iota(\omega_S - p_M \omega_D)$$
$$P_{Mb} = P_{mb} = 0 \tag{C4}$$

at $t = t^*$. To evaluate $E[S^{\rho-1}D]$, we then proceed as follows. The derivative of ψ with respect to ω_D, evaluated at $\omega_D = 0$, is given by

$$\frac{\partial\psi}{\partial\omega_D} = (-\iota)\int\int D\phi(S, D) e^{-\iota\omega_S S} dS\, dD = (-\iota)\int f(S) e^{-\iota\omega_S S} dS, \tag{C5}$$

where

$$f(S) = \int D\phi(S, D)\, dD = E[D \mid S]. \tag{C6}$$

Equation C5 thus shows that the Fourier transform of $E[D \mid S]$ is given by $i(\partial\psi/\partial\omega_D)$. Inversion of this Fourier transform gives

$$E[D \mid S] = \frac{\iota}{2\pi}\int_{-\infty}^{+\infty} \frac{\partial\psi}{\partial\omega_D} e^{i\omega_S S}\, d\omega_S. \tag{C7}$$

We then have (for $\rho < 1$)

$$E[S^{\rho-1}D] = \int_0^\infty S^{\rho-1} f(S)\, dS = \frac{\iota}{2\pi}\int_{-\infty}^{+\infty} \frac{\partial\psi}{\partial\omega_D}\left[\int_0^\infty S^{\rho-1} e^{\iota\omega_S S} dS\right] d\omega_S = \frac{\iota\Gamma(\rho)}{2\pi}\int_{-\infty}^{+\infty} \frac{\partial\psi}{\partial\omega_D}(-\iota\omega_S)^{-\rho}\, d\omega_S. \tag{C8}$$

Combining Equation C8 with Equation 2 leads to an expression for the expected increase of the modifier, in terms of ψ. This allows us to calculate the selection gradient, s_M, acting on the modifier, as defined by (26),

$$s_M = \frac{1}{p_M p_m}\frac{\partial E[\Delta p_M]}{\partial\delta r} = \frac{\iota}{2\pi}N^{-\rho}\Gamma(\rho)\Gamma(1-\rho)\Gamma(1+\rho) \times \int_{-\infty}^{+\infty}\left[\frac{1}{p_M p_m}\frac{\partial^2\psi}{\partial\omega_D\partial\delta r}\right](-\iota\omega_S)^{-\rho}\, d\omega_S, \tag{C9}$$

where the derivative of ψ is evaluated at $\omega_D = 0$, $\delta r = 0$.

The derivative of ψ that appears in Equation C9 can then be expressed in terms of partial derivatives of the P_j's, using Equation 21. Using the fact that P_j variables are of order s_A, one obtains to first order in s_A

$$\frac{1}{p_M p_m}\frac{\partial^2 \psi}{\partial \omega_D \partial \delta r} = -(p_B \tilde{Y}_B + p_b \tilde{Y}_b) \tag{C10}$$

with

$$\tilde{Y}_B = \frac{1}{p_M p_m}\left[p_M \frac{\partial^2 \tilde{P}_{MB}}{\partial \tilde{\omega}_D \partial \delta r} + p_m \frac{\partial^2 \tilde{P}_{mB}}{\partial \tilde{\omega}_D \partial \delta r}\right] \tag{C11}$$

$$\tilde{Y}_b = \frac{1}{p_M p_m}\left[p_M \frac{\partial^2 \tilde{P}_{Mb}}{\partial \tilde{\omega}_D \partial \delta r} + p_m \frac{\partial^2 \tilde{P}_{mb}}{\partial \tilde{\omega}_D \partial \delta r}\right], \tag{C12}$$

where $\tilde{\omega}_D = \omega_D/(2 s_A)$.

Using transformed variables, it can be shown that $\tilde{Y}_B$ and $\tilde{Y}_b$ do not depend on the initial frequency of the modifier, p_M. From Equation 24, one can show that the pairs $(\partial \tilde{P}_{mB}/\partial \tilde{\omega}_D, \partial \tilde{P}_{MB}/\partial \tilde{\omega}_D)$, $(\partial \tilde{P}_{mb}/\partial \tilde{\omega}_D, \partial \tilde{P}_{Mb}/\partial \tilde{\omega}_D)$ remain in the ratio $-p_M/p_m$ for all T (the derivatives being evaluated at $\tilde{\omega}_D = 0$, $\delta r = 0$). We can thus define the quantities

$$\tilde{Z}_{B,\omega_D} = -\frac{1}{p_M}\frac{\partial \tilde{P}_{mB}}{\partial \tilde{\omega}_D} = \frac{1}{p_m}\frac{\partial \tilde{P}_{MB}}{\partial \tilde{\omega}_D}, \quad \tilde{Z}_{b,\omega_D} = -\frac{1}{p_M}\frac{\partial \tilde{P}_{mb}}{\partial \tilde{\omega}_D} = \frac{1}{p_m}\frac{\partial \tilde{P}_{Mb}}{\partial \tilde{\omega}_D}. \tag{C13}$$

We then define

$$\tilde{Z}_{B,\delta r} = \frac{\partial \tilde{P}_{MB}}{\partial \delta r} - \frac{\partial \tilde{P}_{mB}}{\partial \delta r}, \quad \tilde{Z}_{b,\delta r} = \frac{\partial \tilde{P}_{Mb}}{\partial \delta r} - \frac{\partial \tilde{P}_{mb}}{\partial \delta r}. \tag{C14}$$

Using these variables, and after differentiating Equations 24 with respect to $\tilde{\omega}_D$ and δr, one obtains the following differential equations for $\tilde{Y}_B$ and $\tilde{Y}_b$,

$$\frac{\partial \tilde{Y}_B}{\partial T} = [-\theta + (\rho_{AB} - 1)p_b + 2\theta \tilde{P}_B]\tilde{Y}_B - \rho_{AB}p_b(\tilde{Y}_b - \tilde{Z}_{B,\omega_D} + \tilde{Z}_{b,\omega_D}) + 2\theta \tilde{Z}_{B,\omega_D}\tilde{Z}_{B,\delta r} \tag{C15}$$

$$\frac{\partial \tilde{Y}_b}{\partial T} = [-\theta + (\rho_{AB} + 1)p_B + 2\theta \tilde{P}_B]\tilde{Y}_b - \rho_{AB}p_B(\tilde{Y}_B - \tilde{Z}_{b,\omega_D} + \tilde{Z}_{B,\omega_D}) + 2\theta \tilde{Z}_{b,\omega_D}\tilde{Z}_{b,\delta r} \tag{C16}$$

with boundary conditions $(\tilde{Y}_B, \tilde{Y}_b) = (0,0)$ at $T = T^*$. Equations C15 and C16 use the fact that $\tilde{P}_{mB} = \tilde{P}_{MB} = \tilde{P}_B$, and $\tilde{P}_{mb} = \tilde{P}_{Mb} = \tilde{P}_b$ when $\tilde{\omega}_D = 0$ and $\delta r = 0$. Differential equations for the $\tilde{Z}$-variables are also obtained from Equation 24, giving

$$\frac{\partial \tilde{Z}_{B,\omega_D}}{\partial T} = [-\theta + (\rho_{AB} - 1)p_b + \rho_{MA} + 2\theta \tilde{P}_B]\tilde{Z}_{B,\omega_D} - \rho_{AB}p_b\tilde{Z}_{b,\omega_D} \tag{C17}$$

$$\frac{\partial \tilde{Z}_{b,\omega_D}}{\partial T} = [-\theta + (\rho_{AB} + 1)p_B + \rho_{MA} + 2\theta \tilde{P}_B]\tilde{Z}_{b,\omega_D} - \rho_{AB}p_B\tilde{Z}_{B,\omega_D} \tag{C18}$$

with boundary conditions $(\tilde{Z}_{B,\omega_D}, \tilde{Z}_{b,\omega_D}) = (\iota, 0)$ at $T = T^*$, and

$$\frac{\partial \tilde{Z}_{B,\delta r}}{\partial T} = [-\theta + (\rho_{AB} - 1)p_b + \rho_{MA} + 2\theta \tilde{P}_B]\tilde{Z}_{B,\delta r} - \rho_{AB}p_b(\tilde{Z}_{b,\delta r} - \tilde{P}_B + \tilde{P}_b) \tag{C19}$$

$$\frac{\partial \tilde{Z}_{b,\delta r}}{\partial T} = [-\theta + (\rho_{AB} + 1)p_B + \rho_{MA} + 2\theta \tilde{P}_B]\tilde{Z}_{b,\delta r} - \rho_{AB}p_B(\tilde{Z}_{B,\delta r} - \tilde{P}_b + \tilde{P}_B) \tag{C20}$$

with boundary conditions $(\tilde{Z}_{B,\delta r}, \tilde{Z}_{b,\delta r}) = (0,0)$ at $T = T^*$. Finally, differential equations for $\tilde{P}_B$ and $\tilde{P}_b$ are given by

$$\frac{\partial \tilde{P}_b}{\partial T} = -(\theta - p_B)\tilde{P}_b + \rho_{AB}p_B(\tilde{P}_b - \tilde{P}_B) + 2\theta \tilde{P}_b^2 \tag{C21}$$

$$\frac{\partial \tilde{P}_B}{\partial T} = -(\theta + p_b)\tilde{P}_B + \rho_{AB}p_b(\tilde{P}_B - \tilde{P}_b) + 2\theta \tilde{P}_B^2 \tag{C22}$$

with boundary conditions $(\tilde{P}_B, \tilde{P}_b) = (\iota\tilde{\omega}_S, 0)$ at $T = T^*$, with $\tilde{\omega}_S = \omega_S/(2 s_A)$.

In practice, to evaluate the strength of selection on the modifier (s_M) when allele A appears in the population at a given time T, we solve numerically the system of differential Equations C15–C22 for a large array of values of $\tilde{\omega}_S$. The selection gradient s_M is then given by (from Equation C9)

$$s_M = (2s_A)(2Ns_A)^{-\rho}\Gamma(\rho)\Gamma(1-\rho)\Gamma(1+\rho) \times \frac{-\iota}{2\pi}\int_{-\infty}^{+\infty}(p_B\tilde{Y}_B + p_b\tilde{Y}_b)(-\iota\tilde{\omega}_S)^{-\rho}d\tilde{\omega}_S, \tag{C23}$$

where the integral is calculated using the NIntegrateInterpolatingFunction routine of Mathematica (notebook available on request). Finally, it is important to note that as Equations C15–C22 do not depend on the initial frequency of the modifier (p_M), s_M is also independent of p_M.

In the case where loci are in order A–M–B, Equations C17–C20 are changed to

$$\frac{\partial \tilde{Z}_{B,\omega_D}}{\partial T} = [-\theta - p_b + \rho_{MA} + 2\theta\tilde{P}_B]\tilde{Z}_{B,\omega_D} \tag{C24}$$

$$\frac{\partial \tilde{Z}_{b,\omega_D}}{\partial T} = [-\theta + p_B + \rho_{MA} + 2\theta\tilde{P}_b]\tilde{Z}_{b,\omega_D} \tag{C25}$$

$$\frac{\partial \tilde{Z}_{B,\delta r}}{\partial T} = [-\theta - p_b + \rho_{MA} + 2\theta\tilde{P}_B]\tilde{Z}_{B,\delta r} + \rho_{AB}p_b(\tilde{P}_B - \tilde{P}_b) \tag{C26}$$

$$\frac{\partial \tilde{Z}_{b,\delta r}}{\partial T} = [-\theta + p_B + \rho_{MA} + 2\theta\tilde{P}_b]\tilde{Z}_{b,\delta r} + \rho_{AB}p_B(\tilde{P}_b + \tilde{P}_B), \tag{C27}$$

while in the case where loci are in order M–B–A, they become

$$\frac{\partial \tilde{Z}_{B,\omega_D}}{\partial T} = [-\theta + (\rho_{MB} - 1)p_b + \rho_{MA} + 2\theta\tilde{P}_B]\tilde{Z}_{B,\omega_D} - \rho_{MB}p_b\tilde{Z}_{b,\omega_D} \tag{C28}$$

$$\frac{\partial \tilde{Z}_{b,\omega_D}}{\partial T} = [-\theta + (\rho_{MB} + 1)p_B + \rho_{MA} + 2\theta\tilde{P}_b]\tilde{Z}_{b,\omega_D} - \rho_{MB}p_B\tilde{Z}_{B,\omega_D} \tag{C29}$$

$$\frac{\partial \tilde{Z}_{B,\delta r}}{\partial T} = [-\theta + (\rho_{MB} - 1)p_b + \rho_{MA} + 2\theta\tilde{P}_B]\tilde{Z}_{B,\delta r} - \rho_{MB}p_b\tilde{Z}_{b,\delta r} + \rho_{AB}p_b(\tilde{P}_B - \tilde{P}_b) \tag{C30}$$

$$\frac{\partial \tilde{Z}_{b,\delta r}}{\partial T} = [-\theta + (\rho_{MB} + 1)p_B + \rho_{MA} + 2\theta\tilde{P}_b]\tilde{Z}_{b,\delta r} - \rho_{MB}p_B\tilde{Z}_{B,\delta r} + \rho_{AB}p_B(\tilde{P}_b - \tilde{P}_B) \tag{C31}$$

with $\rho_{MB} = r_{MB}/s_B$. Differential equations on $\tilde{Y}_B$, $\tilde{Y}_b$, $\tilde{P}_B$, $\tilde{P}_b$ remain the same.