Theoretical Population Biology **53**, 199–215 (1998)
Article No. TP971358

TPB

Waiting with and without Recombination: The Time to Production of a Double Mutant*

Freddy B. Christiansen

Department of Ecology and Genetics, University of Aarhus, DK-8000 Århus C, Denmark

Sarah P. Otto

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

Aviv Bergman

Center for Computational Genetics and Biological Modeling, Department of Biological Sciences, Stanford University, Stanford, California 94305

and

Marcus W. Feldman

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

Received December 5, 1996

R. A. Fisher and H. J. Muller argued in the 1930s that a major evolutionary advantage of recombination is that it allows favorable mutations to be combined within an individual even when they first appear in different individuals. This effect is evaluated in a two-locus, two-allele model by calculating the average waiting time until a new genotypic combination first appears in a haploid population. Three approximations are developed and compared with Monte Carlo simulations of the Wright–Fisher process of random genetic drift in a finite population. First, a crude method, based on the deterministic accumulation of single mutants, produces a waiting time of $1/\sqrt{N\mu^2}$ with no recombination and $1/\sqrt[3]{\frac{1}{3}R N\mu^2}$ with recombination between the two loci, where μ is the mutation rate, N is the haploid population size, and R is the recombination rate. Second, the waiting time is calculated as the expected value of a heterogeneous geometric distribution obtained from a branching process approximation. This gives accurate estimates for $N\mu$ large. The estimates for small values of $N\mu$ are considerably lower than the simulated values. Finally, diffusion analysis of the Wright–Fisher process provides accurate estimates for $N\mu$ small, and the time scales of the diffusion process show a difference between $R=0$ and for $R \gg 0$ of the same order of magnitude as seen in the deterministic analysis. In the absence of recombination, accurate approximations to the waiting time are obtained by using the branching process for high $N\mu$ and the diffusion

* Research supported in part by Grants 11-9639-1 and 94-0163-1 from the Danish Natural Science Research Council, by a grant from the Research Foundation of Aarhus University, by a grant from the Natural Sciences and Engineering Research Council of Canada, and by NIH Grant 28016.

0040-5809/98 $25.00

approximation for low $N\mu$. For low $N\mu$ the waiting time is well approximated by $1/\sqrt{8N^2\mu^3}$. With $R \gg 0$, the following dependence on $N\mu$ is observed: For $N\mu > 1$ the waiting time is virtually independent of recombination and is well described by the branching process approximation. For $N\mu \approx 1$ the waiting time is well described by a simplified diffusion approximation that assumes symmetry in the frequencies of single mutants. For $N\mu \ll 1$ the waiting time is well described by the diffusion approximation allowing asymmetry in the frequencies of single mutants. Recombination lowers the waiting time until a new genotypic combination first appears, but the effect is small compared to that of the mutation rate and population size. For large $N\mu$, recombination has a negligible effect, and its effect is strongest for small $N\mu$, in which case the waiting time approaches a fixed fraction of the waiting time for $R = 0$. Free recombination lowers the waiting time to about 45% of the waiting time for absolute linkage for small $N\mu$. Selection has little effect on the importance of recombination in general.

1. INTRODUCTION

Arguments concerning the evolutionary advantage of recombination may be classified into two kinds (Felsenstein, 1974; Feldman *et al*., 1997). One of these is developed entirely in terms of selection at the level of the individual and, since the work of Nei (1967) and Feldman (1972), has been couched in terms of modifier genes. The success or failure of recombination is assessed in terms of the fate of alleles of a gene that controls recombination. The mathematical theory for the dynamics of such alleles has yielded the Reduction Principle) which states that if (1) the population is large, (2) mating is random, (3) selection is at the level of viabilities and is constant over time, (4) the modifier locus is neutral with respect to the viabilities acting on the selected loci, (5) mutation does not occur at the selected loci, and (6) a new allele at the modifier locus is introduced near a stable polymorphic equilibrium at which the genes under selection are in linkage disequilibrium, then this new allele will invade the population if it reduces the recombination among the selected loci and it will be expelled if it increases this recombination rate (Feldman *et al*., 1980; Feldman and Liberman, 1986; Zhivotovsky *et al*., 1994).

When some of the conditions under which the Reduction Principle has been proven are violated, the principle is no longer valid. Thus, if there is inbreeding, if selection is at the level of differential fertilities, if viability selection changes over time in a cyclic manner, if selection is strongly directional, if there is segregation distortion, or if the initial equilibrium is due to mutation-selection balance, then the Reduction Principle may fail. That is, numerical and in a few cases analytical examples have been exhibited where a recombination-increasing allele may succeed. When the population is finite and the loci are subject to mutation, the fate of a recombination-modifying allele appears to depend on the viability regime. Studies of recombination modifiers in populations subject to random genetic drift have usually been carried out using numerical simulation and it remains to be determined which classes of assumptions on the selection regime, the mutation process, and the mating system promote the increase of high-recombination alleles (Feldman *et al*., 1997).

The initiation and development of the modifier theory for the evolution of recombination occurred more than 30 years after Fisher (1930b) and Muller (1932) speculated on the evolutionary advantage of recombination. The Fisher–Muller theory attributes the major evolutionary advantage of recombination to the fact that it allows the incorporation into a single individual of advantageous mutations that first appear separately in different individuals. If this were the main reason for the evolution of recombination mechanisms, then "their evolution must have depended on interpopulation rather than intrapopulation selection" (Bodmer, 1970).

Crow and Kimura (1965) extended Muller's original argument by focusing on the rate of incorporation of favorable mutations and showing that this rate was greater in sexual than asexual populations. This was interpreted as the advantage of recombination. Clearly this is a group selection argument. Maynard Smith (1968) replied by demonstrating that in a population with multiplicative selection on advantageous mutations at two loci, linkage equilibrium would be maintained during their segregation and recombination would have no effect on their dynamics. It could not, therefore, be regarded as either advantageous or disadvantageous.

Maynard Smith's argument was developed further by Eshel and Feldman (1970) under the same conditions, but with the double mutant fitness different from the

product of the fitnesses of the advantageous single mutants. If the fitness of the double mutant was higher than the product (supermultiplicative fitness), then a population with recombination would actually have fewer of these advantageous double mutants than one in which recombination is absent. With advantageous double mutants but a submultiplicative fitness, there are more of the double mutants with recombination than without. Recombination therefore can be judged advantageous in the submultiplicative case and disadvantageous in the supermultiplicative case if the group-selection criterion is redefined to be assessed in terms of the frequency of a chromosome.

Bodmer (1970) drew attention to another criterion, namely waiting time, the subject of the present paper. Consider two loci where mutations to an advantageous allele occur at rate μ, and assume that initially at each locus single mutants are present at a low frequency. Bodmer asked how long it takes for the first double mutant to appear and produced an estimate by using a deterministic argument. That is, he estimated the waiting time as the time for the frequency of the double mutant to reach $1/N$ in a population of N haplotypes. His result was that when the recombination between the genes is close to free, then the time until the first appearance of the double mutant in a population with recombination is significantly reduced compared to that in a population without recombination. The reduction was to less than half when $z > 8\mu$, where z is the initial frequency of each of the single mutants. All other things being equal, the presence of recombination must accelerate the first appearance of a double mutant that is initially absent. The subsequent dynamics of the frequencies of these double mutants are, however, determined by the interaction of selection and recombination, as analyzed, for example, in Eshel and Feldman (1970).

Bodmer's (1970) analysis was made in the absence of stochastic effects due to finite population size, but Karlin (1973) included these in his analysis of the problem. He examined the time until production of the first double mutant in the context of a Wright–Fisher model of random genetic drift without selection and showed that its expectation increased with increasing recombination. Karlin demonstrated this mathematically in a population of size $N = 2$ and confirmed the result for larger population sizes using numerical analysis. In each case the mutants were initially absent from the population. The time until fixation of this chromosome might reasonably be considered to be a more appropriate criterion for the advantage of recombination (Karlin, 1973). In our recent analysis of a multi-locus model (Otto *et al.*, 1994), we have shown that under certain selection regimes, the first appearance of the fittest chromosome, as well as its fixation, may be delayed by the presence of recombination and that this result is critically dependent on the initial allele frequencies and on the shape of the selection function.

The time scale of peak shifts in Wright's (1931, 1977) shifting balance process of evolution depends on the rate of spread through a meta-population of an advantageous combination of mutations, each of which is deleterious when by itself. The feasibility of the process depends on the properties of fixation of the advantageous combination within a single deme. Phillips (1996) stressed that the time spent waiting for new mutations dominates the waiting time until fixation in the whole population and, therefore, the time scale of the shifting balance process. Michalakis and Slatkin (1996) argue that for recombination to hasten the fixation of multiple mutants, selection against the single mutants must be weak. Under these conditions, the time until the first appearance of the multiple mutant assumes an important role, since the time until first appearance will be closely related to the time to fixation within the deme if the multiple mutant has a much higher fitness. The dynamics of weakly deleterious single mutants are closely related to the dynamics of neutral or weakly favored mutants, and so analysis of Wright's shifting balance process should take account of the waiting time problems considered by Bodmer (1970) and Karlin (1973). Waiting times obtained under neutral assumptions set upper bounds for waiting times in situations where intermediate forms are advantageous and lower bounds when the intermediate forms are deleterious. The waiting time until the fixation of the advantageous combination of mutations may, however, be prolonged considerably due to the action of recombination when intermediates are deleterious, unless the combination is sufficiently advantageous. This occurs even in two-locus systems with deleterious single mutants, because a stable equilibrium may be produced by the balance between recombination, which erodes the double mutant on the one hand, and the joint action of mutation and selection, which interact to increase the frequency of the favored combination on the other (Eshel and Feldman, 1970; Karlin and McGregor, 1971). We do not develop the analysis of deleterious intermediates any further, and our analysis will focus on two-locus models where single-mutant types are neutral or weakly favored.

The time until the appearance of the first (multilocus) chromosome of a specific desired genotype has relevance in the field of genetic algorithms for adaptive computation. These algorithms write the instructions of a program as a string of 1's and 0's, where each string is seen as a chromosome analog. Rules that mimic the

biological operations of mutation (from 0 to 1 and/or 1 to 0) and recombination among pairs of bitstrings are introduced and the program evolves according to improvement in some criterion of performance. This criterion is usually extremely difficult, if not impossible, to express explicitly in terms of genotypic fitness. The success of a genetic algorithm may be measured by the time it takes to find a useful solution to a problem.

Algorithms of this kind, first introduced by Holland (1975, 1992), have been successful in improving search or sorting routines as well as in solutions of some engineering design problems. The success of genetic algorithms is widely believed to be due to the inclusion of recombination in the genetic algorithm. Little formal theory supports these empirical findings, although it is an active area of research in adaptive computation to circumscribe the class of problems for which solutions are more rapidly obtained with the inclusion of recombination. Solution of a problem is usually couched in terms of the first appearance of a specific bitstring, or a set of bitstrings. This class of computational problems therefore represents a conceptual overlap with the problems posed by Bodmer (1970) and Karlin (1973).

In returning to the two-locus problem of waiting for an advantageous double mutant, we develop in this paper a series of approximations, deterministic and stochastic, in the spirit of Bodmer (1970) and Karlin (1973). We go on to develop a diffusion approximation that uses properties of the Wright–Fisher multinomial sampling process with "killing," following the work of Karlin and Tavaré (1982). These authors sought the time until the first appearance of a recessive phenotype at a single locus. Here we seek the time until the first appearance of a double mutant, with the process being killed when the double mutant appears. Our analysis produces a two-dimensional diffusion which we study by numerical and analytical approximations. We find that estimates of the time to first appearance of a double mutant have different orders of magnitude depending on whether or not there is recombination. This difference is intrinsic to the underlying stochastic process and occurs regardless of the approach taken to estimate the waiting time.

The evaluation of the influence of recombination on the time to production of an advantageous double mutant is a hard analytical problem, and the long debate on the issue should be viewed in this context. The various approaches we develop illustrates the various historical attempts to develop biologically founded simplifications and approximations. Currently, the problem has become more accessible due to an increased ability to perform numerical calculations and simulations. The numerical simulations will be discussed in the light of the various biologically founded simplifications made by earlier authors in order to reach a simple biological description of the effect of recombination on the production of double mutants.

Our analysis addresses properties of the waiting time irrespective of whether or not this is a legitimate criterion upon which to base a discussion of group selection. We shall not enter here into the contentious realm of the validity of the group selection approach.

2. DETERMINISTIC APPROXIMATION

Consider a finite haploid population of size N. At breeding, the population produces an effectively infinite number of gametes which unite at random to produce diploid zygotes. These immediately undergo meiosis to produce haploid offspring, and the offspring population is formed by choosing N individuals at random among these offspring. Thus, the reproduction of the population is according to the Wright–Fisher model (Fisher, 1930b; Wright, 1931; Ewens, 1979). The population is initially monomorphic at two loci, both subject to recurrent mutations between the resident alleles (a and b) and the new alleles (A and B) at the same rate μ. We ignore back-mutations. At meiosis, recombination between the two loci occurs with probability R.

The four genotypes AB, Ab, aB, and ab have fitnesses w, v, v and 1, respectively; that is fitness is measured relative to the resident genotype ab. We are concerned with advantageous or neutral mutants, so we assume $w \geqslant v \geqslant 1$. The frequencies of the four genotypes in the population before breeding are x_1, x_2, x_3 and x_4 (Table 1), and after breeding and selection the expected genotypic frequencies are

$$\begin{aligned} Vx_1' &= w[(x_1 - RD) + \mu(x_2 + RD) \\ &\qquad + \mu(x_3 + RD) + \mu^2(x_4 - RD)], \\ Vx_2' &= v[(1-\mu)(x_2 + RD) + \mu(1-\mu)(x_4 - RD)], \\ Vx_3' &= v[(1-\mu)(x_3 + RD) + \mu(1-\mu)(x_4 - RD)], \\ Vx_4' &= (1-\mu)^2 (x_4 - RD), \end{aligned} \tag{1}$$

where $D = x_1x_4 - x_2x_3$ is the linkage disequilibrium between the loci, and V is the average fitness in the population, namely the sum of the right sides of Eqs. (1). We study the evolution of these frequencies from the initial state where $x_1 = x_2 = x_3 = 0$ and $x_4 = 1$.

Assume initially that the population size is very large to that x_1', x_2', x_3', and x_4' may be considered as the genotypic frequencies before breeding in the offspring

TABLE 1

Population Frequencies and Fitnesses of the Genotypes

Gamete	*AB*	*Ab*	*aB*	*ab*	Σ
Frequency	x_1	x_2	x_3	x_4	1
Fitness	w	v	v	1	

generation. Thus, we consider Eqs. (1) as deterministic recurrence equations. By the symmetry of these equations and the symmetry of the initial state, the two single mutant types *Ab* and *aB* will occur in equal frequencies in the offspring generation and in all subsequent generations, i.e., $x_2^{(n)} = x_3^{(n)}$. In the initial generations the frequency x_2 will be of order μ, so we have

$$x_2' = v(\mu + x_2) + O(\mu^2), \tag{2}$$

where the equation has a leading term of order μ and an error of order μ^2, which we write here as $O(\mu^2)$. Thus, starting from $x_2 = x_3 = 0$ and $x_4 = 1$, we have

$$x_2^{(n)} = x_3^{(n)} \approx \begin{cases} v\mu \dfrac{v^n - 1}{v - 1} & \text{for} \quad v > 1, \\ n\mu & \text{for} \quad v = 1, \end{cases} \tag{3}$$

as long as the frequencies of the mutant genotypes are negligible compared to the frequency of the original chromosome *ab*. This assumption is valid for an extended period only if selection is weak, i. e., when $v = 1 + s$ for s small. For weak selection, (3) becomes

$$x_2^{(n)} = x_3^{(n)} = \mu[n + \tfrac{1}{2}sn(n-1)] + O(\mu^2 n) + O(\mu s^2 n^3). \tag{4}$$

In the deterministic model, the double mutant type is produced immediately at a frequency of order μ^2. However, we consider the deterministic model as an approximation to a process in a large finite population, where the expected number of double mutants is of order $N\mu^2$, which will usually be a small number. Assume, therefore, that *AB* did not occur in the population in generations 0, 1, ..., $n-1$, and so $x_1^{(n-1)} = 0$. Then the expected frequency of double mutant individuals before selection at generation n is

$$\tilde{x}_1^{(n)} \approx \mu^2\{[(2n-1) + s(n-1)(n-2)] + R[(n-1)^2 + s(n-1)^2(n-2)]\}, \tag{5}$$

where the frequencies of single mutants in generation $n-1$ are approximated using Eq. (4). The population is censused before selection so that w, the fitness of double mutants, exerts no influence on the time until the double mutant first appears.

The probability of forming gamete *AB* at or before generation n is therefore approximately given by

$$\sum_{i=0}^{n} \tilde{x}_1^{(i)} \approx \mu^2\{[n^2 + \tfrac{1}{3}sn(n-1)(n-2)] + \tfrac{1}{6}R[(2n-1)n(n-1) + \tfrac{1}{2}s(2n-1)n(n-1)(n-2)]\}. \tag{6}$$

This cumulative probability of occurrence will grow to a level where the occurrence of an *AB* becomes almost certain, and, as a rough indication of the waiting time to the first occurrence of the double mutant gamete, we use the number of generations, T, it takes for the expected number of *AB*-individuals to increase to one. That is, T solves the equation

$$N \sum_{i=0}^{T} \tilde{x}_1^{(i)} = 1. \tag{7}$$

This equation for the waiting time is very closely related to that used by Bodmer (1970). He iterated Eqs. (1) and estimated the waiting time as the solution to $\tilde{x}_1^{(T)} = 1/N$, but the difference between the cumulative frequency in (7) and $\tilde{x}_1^{(T)}$ is of the order of magnitude of the error allowed in the approximation.

Equation (7) can be solved numerically for arbitrary values of μ, N, R, and s. A rough approximation for the waiting time T may be obtained, however, by considering only leading terms in (6). For $s = 0$ and $R = 0$, the dominant term in (6) is $\mu^2 n^2$, and we then find that

$$T_\mu \approx \frac{1}{\sqrt{\mu^2 N}}, \tag{8}$$

which is the same order of magnitude as the expected waiting time found by Karlin (1973). If R is appreciable, however, then the dominant term becomes $\frac{1}{3}R\mu^2 n^3$, giving the waiting time

$$T_R \approx \frac{1}{\sqrt[3]{\frac{1}{3}RN\mu^2}}. \tag{9}$$

In either case, the waiting time is dependent on $N\mu^2$, the expected number of double mutants per generation in the population. The production of an AB individual requires either a double hit (a mutation in an individual that already carries a mutation) or a collision (two individuals that each carry a mutation fuse and recombine), and the probability of either, if selection is weak, is of order μ^2. However, the waiting time for mutation alone is inversely proportional to the square root of $N\mu^2$, whereas the waiting time with significant recombination is inversely proportional to the cube root of $N\mu^2$. The waiting time is therefore of a different order of magnitude with and without recombination, and the differences will be largest whenever $N\mu^2$ is small. We will show that the difference in waiting time with and without recombination is intrinsic to the process and not an artifact of the crude deterministic calculations used here.

3. BRANCHING PROCESS APPROXIMATION

When new mutations initially appear in a population, the number of mutants should be considered as finite even in an infinite population. Thus the dynamics of newly arisen mutants form a stochastic process. Models that take this into account have been considered for one locus by Haldane (1927) and Fisher (1930a). The basis of these models is that in a very large panmictic population, the mutant alleles are rare enough that they exert little influence on one another and may be considered independently. The survival and proliferation of each one is independent of the frequency of mutants, and the number of descendants of a single mutant may be described by a branching process.

Let p_k, $k = 0, 1, 2, ...$, be the probability that an Ab (or aB) individual has k offspring; $p_0 + p_1 + p_2 + \cdots = 1$. The mean number of offspring per mutant individual is λ relative to one for an individual of type ab, i.e., $\sum_{k=0}^{\infty} kp_k = \lambda$. While mutant individuals are rare, the numbers of mutations from a to A and from b to B in the population are assumed each to be Poisson distributed with mean $\theta = N\mu$. At any time, let q_k be the probability that k A alleles are found in the population, and due to the symmetry of the model, the probability of finding k individuals carrying the B allele is also q_k. Initially we have $q_0 = 1$ and $q_1 = q_2 = \cdots = 0$ corresponding to a population of only ab individuals.

We assume the simplest offspring distribution viz. the Poisson distribution, $p_k = \lambda^k/k!\, e^{-\lambda}$. Sums of independent Poisson variables are Poisson distributed, so the probability that i parents of type Ab produce k offspring of type Ab is Poisson distributed with mean $i\lambda$. Thus, the number of A alleles in the offspring population is Poisson distributed with mean $i\lambda + \theta$ in a population with i A alleles ($i \ll N$) among parents, and the recursion equation becomes

$$q'_k = \sum_{i=0}^{\infty} q_i \frac{(i\lambda + \theta)^k}{k!} e^{-(i\lambda + \theta)}. \tag{10}$$

This produces easily iterated recursion equations for the mean and variance of the number of single mutants in the population. The recursion equation for the mean is approximately (2) under weak selection.

The probability of first occurrence of AB at the nth generation is the probability of producing AB in generation n and not in any generation prior to n,

$$d^{(n)} \prod_{i=1}^{n-1} (1 - d^{(i)}), \tag{11}$$

where $d^{(k)}$ is the probability that a double mutant is produced in generation k, $k = 1, 2, ...$, given that it did not occur before. The waiting time to the first occurrence of a double mutant therefore has a heterogeneous geometric distribution with probability parameters $d^{(1)}, d^{(2)}, ...$ An easy parallel to the simple geometric distribution produces the average time to the first occurrence as

$$T_\mu = \sum_{n=1}^{\infty} \prod_{i=1}^{n-1} (1 - d^{(i)}). \tag{12}$$

We will evaluate this waiting time by approximating the probability $d^{(n)}$ using Eq. (10).

3.1. Waiting Time to First Double Mutant, R = 0

The probability $1 - d^{(n)}$ that no AB individual is formed by mutation in any of the Ab individuals in the previous generation is

$$\sum_{k=0}^{\infty} q_k^{(n-1)} (1 - \mu)^k,$$

and the probability that a double mutant is not produced by any aB individuals in the population is the same. Simultaneous mutation of both loci in ab individuals occurs with probability μ^2, so the probability that a double mutant is not produced by any ab individuals in

the population is $e^{-\mu^2 N}$ (Poisson distributed). The above branching process assumptions then yield

$$1 - d^{(n)} = e^{-\mu^2 N} \left(\sum_{i=0}^{\infty} q_i^{(n-1)} (1-\mu)^i \right)^2 \qquad (13)$$

for $n \geqslant 1$ assuming that no recombination occurs, i.e., $R = 0$. A Taylor expansion assuming a small mutation rate gives an approximation to Eq. (13) where the leading term is based on the expected allele frequencies. The approximation is therefore deterministic, and using similar calculations, the resulting expectation for the waiting time is the same as in the deterministic case.

Instead we construct a probabilistic approximate recurrence equation for $1 - d^{(n)}$ by inserting the recurrence Eq. (10) into Eq. (13). To simplify the calculations, we assume no selection ($\lambda = 1$) and obtain

$$1 - d^{(n)} = e^{-\mu^2 N} \left(\sum_{k=0}^{\infty} q_k^{(n-2)} e^{-(k+\theta)\mu} \right)^2.$$

Using the approximation $e^{-k\mu} \approx (1-\mu)^k$, from (13) we have the recurrence equation

$$1 - d^{(n)} = e^{-\mu^2 N} (1 - d^{(n-1)}) + o(\mu), \qquad (14)$$

which is valid for $n \geqslant 2$. Here $o(h)$ is a term such that $o(h)/h \to 0$ as $h \to 0$. From (13) we have $1 - d^{(1)} = e^{-N\mu^2}$, and so Eq. (14) iterates to

$$1 - d^{(n)} = e^{-N\mu^2 (2n-1)} + o(\mu), \quad n \geqslant 1. \qquad (15)$$

Thus, the waiting time to the first occurrence (12) becomes

$$T_\mu \approx \sum_{n=0}^{\infty} \exp(-n^2 N \mu^2). \qquad (16)$$

The order of magnitude of T_μ may be determined by replacing the summation with an integration. This produces the approximation

$$T_\mu \approx \frac{\sqrt{\frac{1}{4}\pi}}{\sqrt{N\mu^2}}, \qquad (17)$$

which is good as long as $N\mu^2$ is small. The waiting time (16) calculated from the heterogeneous geometric distribution is therefore of the same order of magnitude as the deterministic average waiting time (8), but approximately 11% lower for small $N\mu^2$ (less than about 0.1). For larger $N\mu^2$, the waiting time based on the heterogeneous geometric distribution (16) approaches one generation, because the double mutant occurs almost immediately. The deterministic approximation approaches zero, however, and is useless when $N\mu^2 > 1$.

3.2. Waiting Time to First Recombinant Double Mutant

Almost by definition, the production of AB by recombination cannot be handled within the branching process model. The basic assumption of the branching process is that the descendants of every single mutant are independent. They never meet and therefore they cannot recombine. An approximation for large N similar to that used in the description of the mutation process requires a deterministic approximation to the probability that Ab and aB individuals pair. This returns the arguments back to the deterministic approximation discussed in Section 2.

The deterministic approximation may be improved by calculating the waiting time using the heterogeneous geometric distribution to construct a pseudostochastic estimate of the waiting time. Let $c^{(n)}$ be the probability of producing a recombinant double mutant AB in generation n. The probability that none of the N individuals is AB in a population at generation n is approximately

$$1 - c^{(n)} = (1 - x_1^{(n)})^N \approx \exp(-N x_1^{(n)}).$$

Using Eqs. (12) and (6), the average time to first appearance becomes

$$\begin{aligned} T = \sum_{n=0}^{\infty} \exp(-N\mu^2 \{ & [n^2 + \tfrac{1}{3} sn(n-1)(n-2)] \\ & + \tfrac{1}{6} R[(2n-1)\, n(n-1) \\ & + \tfrac{1}{2} s(3n-1)\, n(n-1)(n-2)] \}). \end{aligned} \qquad (18)$$

Integration in place of the summation again produces an approximate value of this expression. The leading order term in the above exponent in the absence of selection and with recombination is $-\frac{1}{3} RN\mu^2 n^2$, for $N\mu^2$ small, and then

$$T_R \approx \frac{\frac{1}{3}\Gamma(\frac{1}{3})}{\sqrt[3]{\frac{1}{3} RN\mu^2}}, \qquad (19)$$

also 11% lower than the deterministic average waiting time (9) (for $N\mu^2 < 10^{-5}$ or $\mu < 10^{-4}$ in Fig. 1). The two approximations, however, are further apart in the presence of recombination than in its absence. When μ

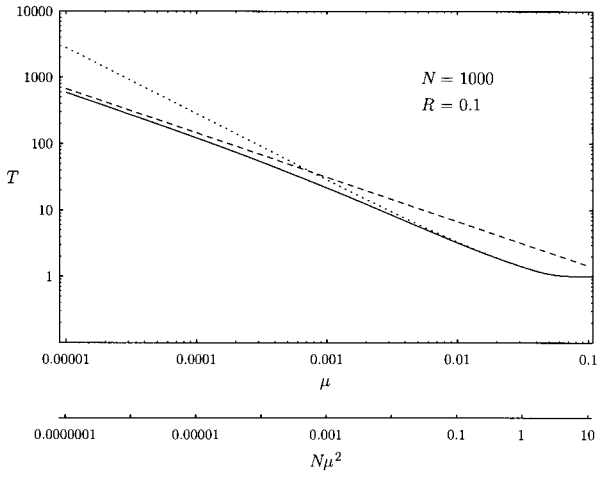

FIG 1. Approximations to the waiting time T until AB first appears with recombination ($R=0.1$) and no selection ($s=0$). The pseudostochastic expected time based on the geometric distribution (18) [solid curve] is compared to the deterministic average waiting time (9) [dashed curved]. For comparison, the expected time based on the branching process approximation (16) for $R=0$ is shown by the dotted curve. Both axes are logarithmic. The abscissa is the mutation rate μ and the ordinate is the waiting time until AB first appears. The second abscissa is $N\mu^2$, the natural parameter of the approximations.

and $N\mu^2$ are high ($\mu \geqslant 0.01$), the estimates based on the geometric distribution with and without recombination are approximately equal (Fig. 1, solid and dotted curves). As the rate of mutation increases, recombination becomes less important in the process of production of AB individuals, and for high mutation rates the process is dominated by mutation. Therefore, we do not expect the deterministic average waiting time T_R in (9) to be a good predictor of the waiting time until AB first appears, since it ignores the possibility that AB is produced by mutation from single mutants. For low mutation rates, however, the order of magnitude difference predicted from the deterministic average waiting times with and without recombination (Eqs. (8) and (9)) is present in the waiting times (16) and (18) based on the heterogeneous geometric distribution (Fig. 1).

The approximations (16) and (19) neglect the effect of possible divergence of the numbers of the two single-mutant types. The frequencies of Ab and aB types are averaged out, and we are left with completely symmetric frequencies in our approximation. For low values of $\theta = N\mu$ this is expected to deviate considerably from the actual situation in a population. For a given total frequency $x_2 + x_3$ of single mutants, the frequency of double heterozygotes, $2x_2x_3$, is maximal for the symmetric situation, where $x_2 = x_3$. Our deterministic approximations therefore overestimate the contribution of recombination with the error increasing as $\theta = N\mu$ decreases.

4. STOCHASTIC ANALYSIS

To allow fluctuations in the relative numbers of single mutants Ab or aB a full stochastic analysis based on the Wright–Fisher model is needed. The offspring population is formed by choosing N individuals at random among newly formed offspring, so that the numbers of the four genotypes in the offspring generation follow a multinomial distribution with frequency parameters x'_1, x'_2, x'_3, and x'_4 given by Eqs. (1). In the initial discussions of the Wright–Fisher model no selection is assumed, i.e., $v = w = 1$.

The numbers of single mutants Ab or aB in the population are i and j, respectively. Before the double mutant AB appears, the remaining $N-i-j$ individuals are ab (Table 2). The stochastic process is killed (we call this state H in the process) when the double mutant AB appears among the N individuals in the population. The expected time to first appearance of AB is, therefore, the expected time for the process to reach state H and be killed. This process is similar to the one analyzed by Karlin and Tavaré (1982).

The probability that one of our N sampled individuals will be a new double mutant is $r_{ij} = \mathrm{P}[AB \mid ij] = x'_1$ from Eqs. (1) for $x_1 = 0$. Similarly, the probabilities that a chosen offspring individual will be of type Ab or aB are $p_{ij} = \mathrm{P}[Ab \mid ij] = x'_2$, and $q_{ij} = \mathrm{P}[aB \mid ij] = x'_3$. The probability of type ab is $s_{ij} = \mathrm{P}[ab \mid ij] = x'_4$. The probability that AB is absent among the offspring, but there are k Ab and ℓ aB individuals in the population in the next generation, is then

$$P_{ij}^{k\ell} = \binom{N}{0 \;\; k \;\; \ell \;\; N-k-\ell} p_{ij}^k q_{ij}^\ell s_{ij}^{N-k-\ell}. \qquad (20)$$

These transition probabilities from state ij to state $k\ell$ have the property

$$\sum_{k=0}^{N} \sum_{\ell=0}^{N-k} P_{ij}^{k\ell} = (1 - r_{ij})^N, \qquad (21)$$

and this is the probability that no AB genotype is sampled. The probability that at least one AB genotype is found in the population is therefore

$$P_{ij}^H = 1 - (1 - r_{ij})^N, \qquad (22)$$

which is the transition probability of the population from state ij to the killed state H. To complete the specification of the transition matrix, we define $P_H^H = 1$ and $P_H^{ij} = 0$. That is, the process, once killed, remains so. Karlin and

TABLE 2

Population Frequencies in the Wright–Fisher Model

Gamete	AB	Ab	aB	ab	Σ
Frequency	x_1	x_2	x_3	x_4	1
Number	0	i	j	$N-i-j$	N
Frequency	0	$\frac{i}{N}$	$\frac{j}{N}$	$1-\frac{i+j}{N}$	1

Taylor (1981) used this process without mutation to address the probability that a recombinant is formed before either locus is fixed.

The Markov process just described may be approximated by a diffusion process using the methods outlined in Karlin and Taylor (1981, chapter 15). The mutation rate is small and the diffusion approximation is obtained by letting $\mu \to 0$ while time goes faster and the population size increases, $N \to \infty$, in such a way that $\theta = N\mu$ stays constant. Essentially, the process is transformed by choosing an appropriate time unit, Δt, and state variables, Y and Z, to describe changes in the frequencies of Ab and aB individuals. We count the number of single mutants in units of N^β individuals, thus

$$Y(t) = \frac{i_t}{N^\beta} \quad \text{and} \quad Z(t) = \frac{j_t}{N^\beta}, \tag{23}$$

where i_t and j_t are the numbers of single mutants present at time t. Time is measured in units of N^α generations, using the unit Δt, so that $N^{-\alpha}\Delta t$ is a single generation. We need to find appropriate values of the scaling parameters α and β to allow the Markov process to converge to a diffusion process as $N \to \infty$. The proper convergence of the moments is secured when $\alpha = \beta$, and then the term describing the killing of the process becomes

$$K(y, z) \approx (y+z)\,\theta N^{2\alpha-1} + \theta^2 N^{\alpha-1} + Ryz N^{3\alpha-1},$$

and we need this term to be positive and finite (Karlin and Taylor, 1981). For $R > 0$ this occurs for $\alpha = \beta = \frac{1}{3}$, but for $R = 0$, we need $\alpha = \beta = \frac{1}{2}$. Thus, the diffusion analysis corroborates the observation from the deterministic approximation that for given $\theta = N\mu$ the waiting time to appearance of a double mutant scales with $N^{1/2}$ for mutation only, and with $N^{1/3}$ when recombination is allowed (Eqs. (8) and (9))

The expected time to the production of AB individuals is $N^\alpha T(y, z)$ generations in a population with $N^\alpha y$ individuals of type Ab and $N^\alpha z$ individuals of type aB, where $T(y, z)$ is the solution to

$$\frac{1}{2}\,y\frac{\partial^2 T}{\partial y^2} + \frac{1}{2}\,z\frac{\partial^2 T}{\partial z^2} + \theta\frac{\partial T}{\partial y} + \theta\frac{\partial T}{\partial z} - KT = -1 \tag{24}$$

The killing term, $K(y, z)$, in this equation is given by

$$K(y, z) = \begin{cases} (y+z)\,\theta & \text{for } R = 0\ (\alpha = \frac{1}{2}), \\ Ryz & \text{for } R > 0\ (\alpha = \frac{1}{3}). \end{cases} \tag{25}$$

The boundary conditions for this equation are given by Eqs. (33) and (34) in Appendix A.

A solution to Eq. (24) with the killing term $(y+z)\,\theta$ is the waiting time T_μ in the process with only mutation. The derivation of the diffusion approximation shows that Eq. (24) with recombination neglects the possibility of production of AB by mutation in Ab and aB individuals, and so a solution is the waiting time T_R.

Equation (24) with $R = 0$ is particularly simple because with no recombination all single mutants are equivalent. AB individuals are produced either by mutation in Ab or aB individuals or by double mutation in ab individuals. The latter possibility is neglected in the diffusion approximation, and the waiting time therefore depends only on $i + j$ or $y + z$, and Eq. (24) can be transformed into the one-dimensional equation

$$\frac{1}{2}\,\xi\frac{d^2\tilde{T}}{d\xi^2} + 2\theta\frac{d\tilde{T}}{d\xi} - \theta\xi\tilde{T} = -1, \tag{26}$$

in the variable $\xi = y + z$. This equation may be solved analytically giving the waiting time in units of $N^{1/2}$ generations. The mean time to the first appearance of a double mutant becomes

$$T_\mu = \frac{\Gamma(\frac{1}{2})\,\Gamma(2\theta)}{\Gamma(2\theta+\frac{1}{2})}\,\frac{\sqrt{\theta}}{\sqrt{2}}\,\frac{1}{\sqrt{\mu^2 N}} \tag{27}$$

generations. This waiting time is of the same order of magnitude as the deterministic and branching process approximations for $\theta \approx 1$. For small mutation rates, however, the diffusion approximation gives a larger estimate of the expected waiting time than the branching process approximation, and taking θ small in Eq. (27) we have the approximation

$$T_\mu \approx \frac{1}{2\sqrt{2\theta}}\,\frac{1}{\sqrt{\mu^2 N}} \tag{28}$$

for the waiting time in generations; numerical evaluations show that this is very good for $\theta \leqslant 0.01$.

For large values of θ the production of single mutants in the population should be described well by a deterministic accumulation of mutants. Therefore, the frequencies of Ab and aB should be close to equal, and we may try to approximate the two-dimensional process by a one-dimensional diffusion approximation by assuming $i=j$ or $y=z$ in the case of recombination. This assumption reduces Eq. (24) with $R>0$ to the equation

$$\frac{1}{4}\zeta\frac{d^2\tilde{T}}{d\zeta^2}+\theta\frac{d\tilde{T}}{d\zeta}-R\zeta^2\tilde{T}=-1, \tag{29}$$

in the variable $\zeta=y=z$. The solution to Eq. (29) is similar to the solution to Eq. (26), but gives the mean time to first appearance of a double mutant for $R>0$ in units of $N^{1/3}$ generations. In generations, the waiting time, $\tilde{T}_R$, for the constrained process with $y=z$ becomes

$$\tilde{T}_R=\frac{\frac{4}{3}\Gamma(\frac{1}{3})\,\Gamma(\frac{4}{3}\theta)}{\Gamma(\frac{4}{3}\theta+\frac{2}{3})}\,\frac{\theta^{2/3}}{\sqrt[3]{36}}\,\frac{1}{\sqrt[3]{\frac{1}{3}R\mu^2N}}. \tag{30}$$

This approximation to the waiting time is again of the same order of magnitude as the deterministic and pseudostochastic approximations for $\theta\approx 1$, while for small θ, it produces a larger estimate.

As R tends to zero, the production of double mutants by recombination will decrease, and the waiting time in (30) tends to infinity. Close to the limit as R tends to zero, the behavior of the system is described better by (27), where the effect of recombination is neglected. We may however, extend this description to low values of the recombination frequency by scaling R as $R\approx\omega N^{-1/2}$, where ω is the scaled recombination parameter that remains constant as $N\to\infty$. The waiting time until the first appearance of ab is then the solution to Eq. (24) with the killing term

$$K(y,z)\approx(y+z)\,\theta+yz\omega, \tag{31}$$

$\alpha=\beta=\frac{1}{2}$, and the same boundary conditions as before.

The diffusion process may also be used in the case of weak selection (see Karlin and Tavaré, 1981) as an approximation to the Fisher–Wright model based on the full recursion (1). The diffusion approximation remains valid if we may scale the selection coefficient s by N^α so that $S=sN^\alpha$ is a constant as $N\to\infty$. Equation (24) is then replaced by

$$\frac{1}{2}y\frac{\partial^2T}{\partial y^2}+\frac{1}{2}z\frac{\partial^2T}{\partial z^2}+(yS+\theta)\frac{\partial T}{\partial y}+(zS+\theta)\frac{\partial T}{\partial z}-KT=1, \tag{32}$$

where K is given by (25) and the same boundary conditions apply. Thus, the effect of selection of this strength on the process is small.

5. NUMERICAL ANALYSES

The deterministic and stochastic approximations derived above will be compared to Monte Carlo simulations of the Wright–Fisher process without selection in a population of size $N=1000$. The generation in which an AB individual appears for the first time is recorded and the simulation is repeated 10,000 times to obtain a mean waiting time $\hat{T}$. The results of these simulations are shown in Table 3 and as open circles in Figs. 2, 3 and 4 to be discussed in the following.

5.1. Waiting Times for Absolute Linkage

Figure 2 shows the mean waiting times, $\hat{T}_\mu$, without recombination (open circles) and compares these to the various theoretical estimates. The average waiting time estimated by the diffusion approximation (27) is shown by the dashed curve, and the correspondence between the approximation and the simulation is very good with a clear deviation only for $\theta=N\mu\geqslant 10$. This deviation is caused by the trivial fact that time is assumed to be continuous and not discrete in the diffusion analysis, so that as the mutation rate becomes large the expected time until the double mutant is produced approaches zero in the diffusion process rather than one, which is the minimum waiting time in a process with discrete generations.

The branching process approximation (16), shown by the solid curve in Fig. 2, underestimates the waiting time

TABLE 3

The Effect of Recombination on the Mean Waiting Time in Case of No Selection

	R		
$N\mu$	0	0.1	0.5
0.01	11012 ± 74	7277 ± 68	5094 ± 48
0.03	2293 ± 21	1451 ± 13	1032 ± 10
0.1	447 ± 4	291 ± 2	200 ± 2
0.3	115.2 ± 0.8	81.3 ± 0.5	57.2 ± 0.4
1	30.9 ± 0.2	24.3 ± 0.1	18.2 ± 0.1
3	10.17 ± 0.05	9.07 ± 0.04	7.54 ± 0.03
10	3.36 ± 0.02	3.26 ± 0.01	3.04 ± 0.01
30	1.43 ± 0.01	1.43 ± 0.01	1.43 ± 0.01

Note. Shown are the average waiting time and the standard deviation in 10,000 simulations of the Wright–Fisher model with $N=1000$.

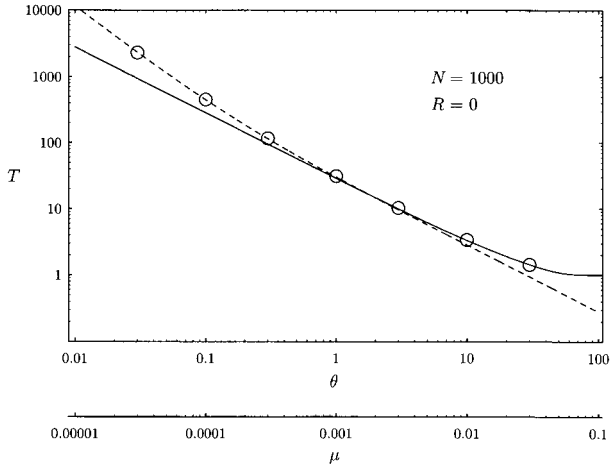

FIG. 2. The waiting time T until AB first appears in the case of absolute linkage ($R=0$) and no selection ($s=0$). The mean waiting time observed in simulations of the Wright–Fisher model [open cycles] is compared to the diffusion waiting time (27) [dashed curve] and the expected time based on the branching process approximation (16) [solid curve]. Both axes are logarithmic. The abscissa is $\theta = N\mu$ and the ordinate is the waiting time until AB first appears (for comparison a second abscissa showing the mutation rate μ is given).

for small θ, but provides the best estimate for very large θ. The underestimation of the waiting time for small θ is substantial. For $\theta = 0.03$ ($\mu = 3 \times 10^{-5}$), the branching process estimate of the waiting time T_μ is 935 while the diffusion estimate is more than double, namely 2326, which is very close to the simulated value of 2293 generations (Table 3).

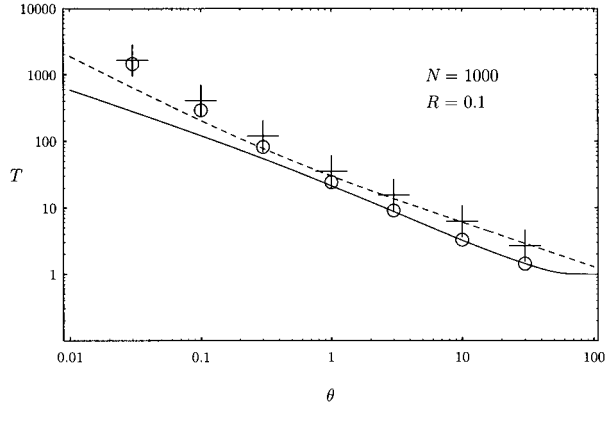

FIG. 3. The correspondence, for $R=0.1$ and no selection, between the various estimates of the waiting time T until AB first appears. The mean waiting time observed in simulations of the Wright–Fisher model [open circles] is compared to the numerical solution to the two-dimensional diffusion (24) [crosses], the symmetric diffusion waiting time (30) [dashed curve], and the pseudostochastic expected time based on the geometric distribution (18) [solid curve] (from Fig. 1). Axes are as in Fig. 2.

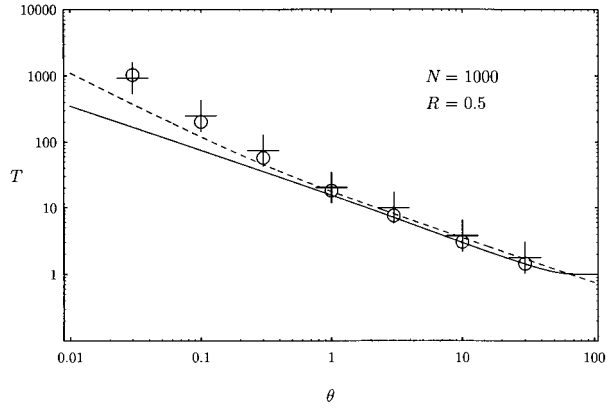

FIG. 4. The correspondence, for $R=0.5$ and no selection, between the various estimates of the waiting time T until AB first appears. Symbols, curves and axes are as in Fig. 3.

The analysis that leads to the deterministic and branching process approximations of the waiting time assumes that the flux of mutants into the population is constant. This amounts to assuming that with high probability the first double mutant is produced while the single mutants are still rare in the population. This biological model therefore provides a good approximation of the process in a large population ($\theta > 1$). In a small population, however, random genetic drift may carry the single mutants to high frequencies, and the resulting appreciable decrease in the number of ab individuals results in a lowered influx of mutants in the population. This effect is included in the diffusion approximation of the process, and it becomes noticeable for $\theta < 1$. The diffusion approximation of the waiting time, on the other hand, becomes inaccurate for $\theta > 1$.

5.2. Waiting Times with Recombination

Figures 3 and 4 show the mean waiting time, $\hat{T}$, for recombination rates $R=0.1$ and $R=0.5$, respectively. The approximation (18) of T based on the pseudostochastic, branching process estimation (solid curve) substantially underestimates the waiting time for small θ, and provides the best estimate for very large θ. This is the same pattern as in the case of absolute linkage (Fig. 2). The high frequencies of single mutants caused by random genetic drift in a finite population may produce the discrepancy for small θ, and the diffusion approximation of T_R is expected to bridge the gap. The underestimation of the waiting time with recombination, however, may have an additional source. Asymmetry in the frequencies of Ab and aB has no effect for absolute linkage because

the only source of AB gametes is then mutation in a single mutant gamete, any single-mutant gamete. Any asymmetry in the single-mutant frequencies, however, will lower the probability of forming an AB gamete by recombination.

To evaluate this effect we compare the waiting time $\tilde{T}_R$ obtained by the symmetric diffusion approximation (30) with the other estimates. In Figs. 3 and 4 the estimate $\tilde{T}_R$ is shown by the curves with long dashes. For small θ's the symmetric diffusion approximation substantially underestimates the mean waiting time in the Monte Carlo simulation. The only expression that gives the expected waiting time, when the frequencies of Ab and aB individuals are allowed to differ, is the general diffusion equation (24). An analytical solution to Eq. (24) is unknown, but a numerical solution may be obtained using a relaxation method and a finite difference approximation (Appendix B). In Figs. 3 and 4, solutions T_R to Eq. (24) are shown by crosses, and we find an excellent correspondence between the mean waiting time $\hat{T}$ in the Monte Carlo simulation and the diffusion approximation for $\theta \leqslant 1$. For $\theta > 1$, the general diffusion approximation and the symmetric diffusion approximation agree, but overestimate the waiting time.

The diffusion approximation neglects the production of AB by mutation in Ab and aB individuals, and this seems reasonable for low mutation rates, $\theta < 1$. For higher mutation rates, however, mutation is a major source of AB individuals, and the waiting time in the Monte Carlo simulations for $\theta = 10$ and $\theta = 30$ is virtually independent of the recombination frequency (the values in Figs. 2, 3 and 4 are indistinguishable; see Table 3). The pseudostochastic approximation based on the geometric distribution (18) gives the best fit, and it depends little on recombination for large θ (Fig. 1). Therefore, for high mutation rates, $\theta > 1$, the mutation process dominates the production of AB individuals and recombination may be neglected. The best description is obtained by using the assumption that Ab and aB individuals are still rare when AB is first produced (the branching process approximation). For low mutation rates, $\theta \leqslant 1$, the production of AB individuals is predominantly by pairing and recombination between Ab and aB individuals, and the best description is obtained by neglecting the production of AB by mutation (the diffusion approximation).

Figures 3 and 4 show that the numerical solutions to the diffusion equation (24) for finite R with the killing term from Eq. (25) provide results in good agreement with the results obtained by simulation of the process. We were unable, however, to obtain satisfactory numerical solutions from the diffusion equation (24) by using the killing term (31) for small R. The equation for small R estimates T and should therefore allow a numerical approximation to T_μ for $R = 0$, but our numerical solutions of the equation overestimate the waiting time for $\theta < 1$ when compared to the analytical solution (27) of the equation. For $\theta > 1$ the waiting times obtained with the small R approximation are very close to the analytical solutions, but these analytical solutions give inaccurate mean waiting times for large values of θ (Fig. 2).

5.3. The Effects of Selection

The mean waiting time, $\hat{T}$, for a one per cent advantage of the single mutant gametes ($s = 0.01$) was determined in simulations of the Wright–Fisher model. The results of these simulations are shown in Table 4 and as closed circles in Figs. 5, 6 and 7. Figure 5 compares the mean waiting time, $\hat{T}_\mu$, without recombination and the pseudostochastic branching process approximation (18), shown by the solid curve. The pattern of deviations of the approximate value from the simulated value is very similar to that seen in the case of no selection (Fig. 2), and for comparison open circles and the dashed curve in Fig. 5 reproduce the simulated waiting times and branching process approximation (16) in the case of no selection. We were unable to obtain satisfactory numerical solutions to the diffusion equation (32).

The effect of selection is evidently to shorten the time until the first appearance of the double-mutant type, but Fig. 6 clearly shows that the effect is fairly modest, unless mutation rates are very low. The magnitudes of the effect

TABLE 4

The Effect of Recombination on the Mean Waiting Time in Case of Weak Selection Favoring the Single Mutants ($s = 0.01$)

	R		
$N\mu$	0	0.1	0.5
0.01	2703 ± 24	2488 ± 22	2255 ± 20
0.03	938 ± 8	755 ± 6	643 ± 6
0.1	282 ± 2	213 ± 2	163 ± 1
0.3	95.9 ± 0.6	68.6 ± 0.4	50.7 ± 0.3
1	28.9 ± 0.2	23.2 ± 0.1	17.4 ± 0.1
3	9.93 ± 0.05	8.93 ± 0.04	7.34 ± 0.03
10	3.29 ± 0.02	3.23 ± 0.01	3.02 ± 0.01
30	1.43 ± 0.01	1.43 ± 0.01	1.41 ± 0.01

Note. Shown are the average waiting time and the standard deviation in 10,000 simulations of the Wright–Fisher model with $N = 1000$.

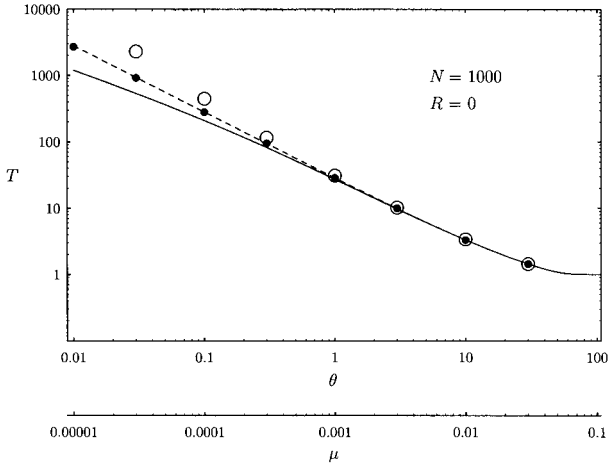

FIG. 5. The waiting time T until AB first appears in the case of absolute linkage ($R = 0$). The expected time based on the geometric distribution (18) is shown for $s = 0.01$ [solid curve] and for no selection ($s = 0$) [dashed curve] (see also Eq. (16)). The mean awaiting time observed in simulations of the Wright–Fisher model is shown by closed circles for $s = 0.01$ and by open circles for no selection (Fig. 2). The simulated values with selection appear to align with the no-selection expected time, but this is coincidental. Axes are as in Fig. 2.

of recombination with and without selection are similar (Fig. 6), and in both cases the effect of recombination is apparent even for quite small values of the recombination frequency. The effect of recombination on the waiting time is modest for a range of mutation rates with the value of selection used (Fig. 7), and the effect of recombination decreases both towards larger and smaller mutation rates with a maximum around $\theta = 1$.

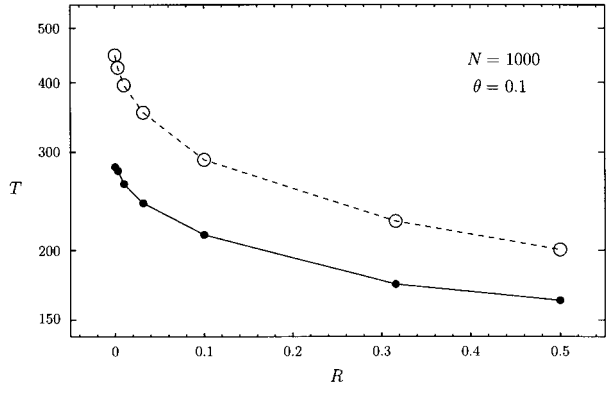

FIG. 6. Comparison of the waiting time T until AB first appears in the cases of selection and no selection. The mutation rate is $\mu = 10^{-4}$. The mean waiting time observed in simulations of the Wright–Fisher model is shown by closed circles for $s = 0.01$ and by open circles for no selection. The abscissa is the recombination frequency R and the ordinate is the waiting time until AB first appears. The ordinate is logarithmic.

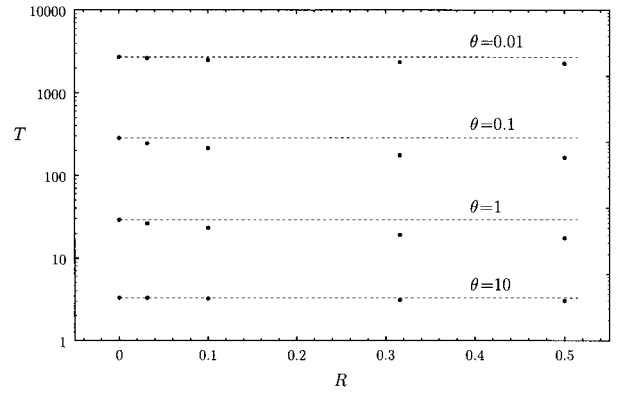

FIG. 7. The effect of selection, recombination and mutation on the waiting time T until AB first appears in a population of size $N = 1000$. The selection coefficient is $s = 0.01$. For each value of $\theta = N\mu$ the mean waiting time observed in simulations of the Wright–Fisher model is shown by closed circles, and the waiting times for $R = 0$ are given by dashed lines for comparison. Axes are as in Fig. 6.

6. CONCLUSIONS

The incorporation of new mutations within a population underlies all longterm evolutionary change. Various aspects of this process may be studied, including the probability of fixation of new mutants, the time to their fixation, and the time until novel genotypic combinations appear within a population. We focus on the last of these, calculating the waiting time until a new genotypic combination first appears in a haploid population using a two-locus, two-allele model. The classical evaluation of this waiting time is obtained by assuming that the variants at both loci are still rare when they occur together in a gamete for the first time. This assumption leads to two approximations. The first approach followed the deterministic accumulation of single mutants, Ab and aB, and estimated the waiting time as the number of generations until the expected total production of double mutants, AB, reached one. This deterministic calculation estimated the time to first appearance of the double-mutant genotype as $1/\sqrt{N\mu^2}$ in the absence of recombination and $1/\sqrt[3]{\frac{1}{3}RN\mu^2}$ in the presence of recombination. The second approach used a branching process approximation to the waiting time. Without recombination this produces the waiting time (16) as the expected value of a heterogeneous geometric distribution. With recombination the branching process approximation necessarily breaks down, but a pseudostochastic approximation to the waiting time is constructed by using the deterministic recursion equations to estimate the probabilities of production of AB at each generation. These are then used as parameters of the heterogeneous

geometric distribution, and the expected waiting time is given by Eq. (18). When compared with Monte Carlo simulations of the Wright–Fisher process in a population of $N = 1000$, this approach provided good estimates for the waiting times when mutation rates are high ($N\mu > 1$; Figs. 2, 3 and 4). For low mutation rates ($\theta = N\mu < 1$), however, the estimates may deviate considerably from the simulation values. Thus, the branching process assumption that mutants are combined while rare provides a good description of the process in a large population. This approximation is in essense a deterministic approximation, as the effect of random genetic drift is neglected. Random genetic drift may carry the population frequency of a mutant to appreciable frequencies, and to investigate the effect of these stochastic fluctuations we used a diffusion analysis of the Wright–Fisher process.

The diffusion approximation for the waiting time to first appearance of the double-mutant genotype in the absence of recombination, Eq. (27), provides better estimates than the branching process approximation when mutation rates are low ($N\mu < 1$), but worse estimates for high mutation rates (Fig. 2). The two approximations for the waiting time make different simplifying assumptions about the process of accumulation of single mutants and about the production of the double-mutant genotype. The branching process assumes that the double-mutant genotype is produced before the single mutants reach appreciable frequencies. The accumulation of mutants occurs roughly at the rate μ and the expected number of double-mutant individuals increases at the rate $N\mu^2$. Thus, when $N\mu$ is large, the assumption should be satisfied, and this is in agreement with the excellent description by the branching process when $N\mu \geqslant 1$ (Fig. 2). The diffusion approximation, as a continuous time approximation, assumes that $N\mu$ is small, and it provides an equally excellent description when $N\mu \leqslant 1$ (Fig. 2). The overlap of the two approximations for $N\mu \approx 1$ is good, and so in combination the two approximations describe the pure mutation process: when $N\mu \geqslant 1$ the waiting time is described by the process of accumulation of mutations in a large population and when $N\mu \leqslant 1$ the process is influenced by random genetic drift due to a limited population size.

General diffusion approximations with recombination behave very much like the approximations without recombination. The numerical solutions of the diffusion equation agree well with the simulation results for low mutation rates ($N\mu < 1$; Figs. 3 and 4), but for high mutation rates ($N\mu > 1$) the pseudostochastic waiting times based on the deterministic description are far superior. The simpler diffusion approximation (30) that assumes symmetric population frequencies at the two loci agrees with the general diffusion approximation except when mutation rates are low ($N\mu < 1$). The correspondence with the simulation results is only usable for intermediate mutation rates ($N\mu \approx 1$). Thus, the assumptions that mutants are rare and that the population frequencies at the loci are symmetric both break down for $N\mu < 1$. For low mutation rates, the process must therefore have at least one of the mutants at an appreciable frequency, very different from the frequency of the mutant at the other locus.

For high mutation rates ($N\mu > 1$), the branching process approximation and the pseudostochastic approximation provide good descriptions of the average waiting time until the first double mutant occurs with recombination. The waiting times with and without recombination are, however, very similar (Fig. 1 and Table 3), and so the branching process without recombination provides a good description of the general process for high mutation rates. The branching process assumes that the double-mutant genotype is produced before the single mutants reach appreciable frequencies, and this is consistent with the observation that recombination is of minor importance. Selection on the rare single mutants is slow, and so our description of the process is consistent with the observation that selection has little effect when mutation rates are high (Fig. 5).

For low mutation rates ($N\mu < 1$), the process is different because of the strong influence of random genetic drift due to a limited population size. The two loci behave asymmetrically and a single mutant can reach high frequencies before the first double mutant is produced. We may use the approximation in Eq. (28) as a natural base for the comparison of the waiting times with and without recombination (Fig. 8). The simulation values of the waiting time without recombination are very close to the approximation for $T_\mu > 1000$ ($\theta = N\mu < 0.1$; Table 3). The influence of recombination on the difference among the simulation values in Fig. 8 increases as the waiting time increases to the order of magnitude $T \approx 100$ ($\theta = 0.3$), but for higher waiting times the ratios between simulation values for given recombination frequencies stays constant. The observed values for $N = 1000$ from simulations give $T_{R=0.1}/T_{R=0} \approx 0.65$ and $T_{R=0.5}/T_{R=0} \approx 0.45$. Thus, for low values of $N\mu$ the waiting time is

$$T_\mu \approx \frac{1}{\sqrt{8\mu^3 N^2}}$$

for no recombination, from Eq. (28), and the effect of recombination is to reduce this waiting time by a

constant fraction; about $\frac{1}{3}$ for $R = 0.1$ and about $\frac{1}{2}$ for $R = 0.5$. Thus, recombination has a definite effect on the waiting time when mutation rates are low, although this effect is weak (Fig. 8). The pseudostochastic and the symmetric diffusion approximations to the waiting time predict that this time is inversely proportional to $\sqrt[3]{R}$. This relationship obviously does not extend to small R, but neither does it extend to small $N\mu$ or large $N\mu$. For large $N\mu$, the waiting time is virtually independent of R. For small $N\mu$, we have $T_{R=0.5}/T_{R=0.1} \approx 0.7$ for the simulated values with $N\mu < 1$ (Table 3) compared to an expectation of $0.58 (= 1/\sqrt[3]{5})$. We see these results for low mutation rates as a reflection of the pronounced asymmetry in the frequencies of single mutants due to random genetic drift in the population.

The waiting time for very low mutation rates ($N\mu < 0.1$) is dominated by the time until the second mutation occurs, whether or not the double mutant may be produced by recombination. The population frequencies of single mutants are asymmetric, and the observed pattern may be explained by similar probabilities for a second mutational hit in the common single mutant and for a mating of the common and the rare single mutants. The qualitative effects of recombination on this process are not expected to be influenced by selection; rather, the asymmetry is expected to be exaggerated by selection. This is in agreement with the observation that with selection, the influence of recombination on the waiting time is considerably weaker for $N\mu = \theta = 0.01$ than for $N\mu = 0.1$ (Fig. 7). The observed values of $T_{R=0.1}/T_{R=0}$ and $T_{R=0.5}/T_{R=0}$ from simulations with selection calculated from Table 4 are all larger than or equal to the observed values without selection from Table 3. Thus, the influence of recombination is weaker with than without selection, and for low mutation rates this difference is very pronounced, e.g. for $N\mu = 0.01$ we observe $T_{R=0.1}/T_{R=0} = 0.92$ and $T_{R=0.5}/T_{R=0} = 0.83$ for $s = 0.01$ compared to 0.66 and 0.46, respectively, for $s = 0$. The significantly shorter waiting time for low mutation rates with selection is therefore mainly due to exaggerated asymmetry in the population frequencies of single mutants making the second mutational hit more likely.

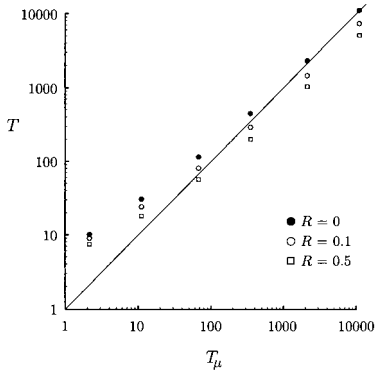

FIG. 8. The waiting time T until AB first appears with and without recombination in a population of size $N = 1000$. For each value of $N\mu$, corresponding values (T_μ, T) are shown, where T_μ is the diffusion approximation (28) for small θ and T is the mean waiting time observed in simulations of the Wright–Fisher model. The values in Table 3 are used for $\theta < 10$ and are shown by closed circles for $R = 0$, open circles for $R = 0.1$ and open squares for $R = 0.5$. The line of equality, namely (T_μ, T_μ), is given for comparison. The axes show the waiting time until AB first appears; T_μ varies along the abscissa and T along the ordinate. Both axes are logarithmic.

Our results indicate that the time to first appearance is most sensitive to the mutation rate and less sensitive to the recombination rate (Fig. 7). The sensitivity to the population size is revealed by the diffusion approximation (27) which is an excellent approximation to the process when recombination is absent and mutation is rare. The population size only influences the waiting time through $\theta = N\mu$, whereas the waiting time has an additional, inverse proportionality to $\sqrt{\mu}$. In a population with $\theta = 0.1$, e.g. of size 1000 with mutation rate 10^{-4}, increasing the population size by a factor of 10 reduces the waiting time by a factor of 4.7, but increasing the mutation rate by a factor of 10 decreases the waiting time by a factor of 14.9.

The influence of recombination on the waiting time to first appearance of a double mutant gamete is very small compared to the expectation formulated by Muller (1932). The maximum effect occurs for low mutation rates, where free recombination lowers the waiting time to just less than half the waiting time found for absolutely linked loci in the case of no selection on the single mutants. Selection makes the waiting time shorter, but the influence of selection on the relative effect of recombination is less pronounced, and the influence seems to vanish as the mutation rate becomes small. Thus with a low mutation rate, recombination has no qualitative effect on the waiting time, but is limited to a minor quantitative influence.

The waiting time to the first occurrence of an advantageous combination of mutants is long when the mutation rate is low. Random genetic drift plays a crucial rôle in the process of waiting. A single mutant typically drifts to appreciable frequencies before a double mutant is produced, and the intuitive description of the process as a combination of rare mutants is wrong. Therefore, we expect the waiting time to the formation of a combination of individually disfavored mutants, a statistic important for shifting balance evolution, to be considerably longer.

APPENDIX A: BOUNDARY CONDITIONS

At the boundaries of the frequency domain for Ab and aB we have $i=0$, $j=0$ or $i+j=N$ (Table 2). For $i+j=N$ every mutation produces AB, and the waiting time for a mutation event is $1/\theta$ generations, which, after scaling, is equivalent to the boundary condition

$$T(y, \infty) = T(\infty, z) = 0 \tag{33}$$

for all $y>0$ and $z>0$. Thus, in scaled time, AB is produced immediately. With recombination, these boundary conditions remain valid in that the waiting time to production of a recombinant AB is about

$$\frac{1}{N \times R \times \frac{i}{N}\left(1-\frac{i}{N}\right)}$$

generations, and therefore, after scaling, AB is produced immediately.

The behavior of T along the boundary $i=0$ or $y=0$ may be deduced from the following argument. For a given z, $0<z<\infty$, the difference between $T(0, z)$ and $T(y, z)$ when y is sufficiently small, is essentially the time to produce $N^{\alpha}y$ mutants in the population. This time is $N^{\alpha}y/(N\mu)$ generations, which, after scaling, gives us

$$T(y, z) - T(0, z) \approx -\frac{y}{\theta}$$

for $y \to 0$. Thus, at the boundaries, we require

$$\frac{\partial T}{\partial y}(0, z) = -\frac{1}{\theta} \quad \text{and} \quad \frac{\partial T}{\partial z}(y, 0) = -\frac{1}{\theta} \tag{34}$$

for all $0<y<\infty$ and $0<z<\infty$. The boundary conditions (33) and (34) are sufficient to secure a unique solution to Eq. (24).

APPENDIX B: NUMERICAL SOLUTION FOR THE DIFFUSION APPROXIMATION

We wish to solve the nonhomogeneous elliptic equation (24) with $K(y, z) = Ryz$ and boundary conditions (33) and (34). That is, we wish to find the function $T(y, z)$ that solves $LT(y, z) = -1$, using the mixed boundary conditions: $T(y, z) = 0$ on the boundaries $y=\infty$ and $z=\infty$ (a Dirichlet boundary condition) and $\partial T/\partial y = \partial T/\partial z = -1/\theta$ on the boundaries $y=0$ and $z=0$ (a Neumann boundary condition). As discussed in section 17.4 of Press *et al.* (1992), the solution may be obtained by finding the equilibrium solution the associated diffusion equation

$$\frac{\partial T_t(y, s)}{\partial t} = LT_t(y, z) + 1. \tag{35}$$

As $T_t(y, z)$ approaches an equilibrium and $\partial T_t(y, z)/\partial t$ approaches zero, $T_t(y, z)$ is said to *relax* from an initial distribution to one that solves (24), i.e., $T_t(y, z) \to T(y, z)$ as $t \to \infty$. The function $T_t(y, z)$ is represented on a finite grid of size $M \times M$ surrounded by a perimeter of width one, with a grid spacing of Δ_y along the y-axis and Δ_z along the z-axis. Each point on the grid is given an initial value, which does not solve $LT(y, z) = -1$. A finite difference method is then used to update each point on the grid over time until $\partial T_t(y, z)/\partial t$ is approximately zero. The values of $T_t(y, z)$ on the grid then form an approximate numerical solution to $LT(y, z) = -1$, denoted by $\check{T}$.

Each point on the grid (j, k), $j,\ k = 1, 2, ..., M$, is updated from one round (numbered n) to the next $(n+1)$ by approximating the differential equation (35) with a finite-difference equation. For internal points, $j, k = 1, 2, ..., M$, we use the equation

$$\begin{aligned}
&\frac{\check{T}^{(n+1)}(j, k) - \check{T}^{(n)}(j, k)}{\Delta t} \\
&\quad = \frac{j\Delta_y}{2} \frac{\check{T}^{(n)}(j-1, k) - 2\check{T}^{(n)}(j, k) + \check{T}^{(n)}(j+1, k)}{\Delta_y^2} \\
&\qquad + \frac{k\Delta_z}{2} \frac{\check{T}^{(n)}(j, k-1) - 2\check{T}^{(n)}(j, k) + \check{T}^{(n)}(j, k+1)}{\Delta_z^2} \\
&\qquad + \theta \frac{\check{T}^{(n)}(j+1, k) - \check{T}^{(n)}(j, k)}{\Delta_y} \\
&\qquad + \theta \frac{\check{T}^{(n)}(j, k+1) - \check{T}^{(n)}(j, k)}{\Delta_z} \\
&\qquad - Rj\,\Delta_y k\,\Delta_z \check{T}^{(n)}(j, k) + 1,
\end{aligned} \tag{36}$$

where Δt is the time step corresponding to one round of relaxation. The perimeter of the grid is not updated by (36), but by the equations

$$\check{T}(M+1, k) = \check{T}(j, M+1) = \check{T}(M, k) = \check{T}(j, M) = 0,$$

which correspond to the boundary conditions (33), and

$$\check{T}(0, k) = \check{T}(1, k) + \frac{\Delta_y}{\theta},$$

$$\check{T}(1, k) = \check{T}(2, k) + \frac{\Delta_y}{\theta},$$

$$\check{T}(j, 0) = \check{T}(j, 1) + \frac{\Delta_z}{\theta},$$

$$\check{T}(j, 1) = \check{T}(j, 2) + \frac{\Delta_z}{\theta},$$

which correspond to the boundary conditions (34). The first internal cell (1, 1) in the grid, represents the point $y = 0$, $z = 0$ and, in general, a point (j, k) not on the boundary of the grid represents the point $y = (j-1)\,\Delta_y$, $z = (k-1)\,\Delta_z$. That is, $\check{T}^{(n)}(j+1, k+1)$ estimates $T(j\,\Delta_y, k\Delta_z)$ after n rounds of performing the relaxation method.

In applying the relaxation method, we used a two dimensional grid with $M = 100$ and a spatial resolution of $\Delta_y = \Delta_z = 0.1$ which places the values of y and z in the interval $[0, 9.9]$. The numerical iteration was stopped when the value of the function at the origin changed from one time step to the next by less than 10^{-8}. Numerical evaluations showed that the chosen parameters were sufficient to obtain convergence and stability of the solution.

REFERENCES

Abramowitz, M., and Segun, I. A. 1970. "Handbook of Mathematical Functions," Dover, New York.

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

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

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

Ewens, W. J. 1979. "Mathematical Population Genetics," Biomathematics, Vol. 9, Springer-Verlag, Berlin/Heidelberg/New York.

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

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

Feldman, M. W., and Liberman, U. 1986. An evolutionary reduction principle for genetic modifiers, *Proc. Natl. Acad. Sci. USA* **83**, 4824–4827.

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

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

Fisher, R. A. 1930a. The distribution of gene ratios for rare mutations, *Proc. Roy. Soc. Edinburgh* **50**, 205–220.

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

Haldane, J. B.S. 1927. A mathematical theory of natural and artificial selection. Part IV. Selection and mutation, *Proc. Cambridge Philos. Soc.* **23**, 838–844.

Holland, J. H. 1975. "Adaptation in Natural and Artificial Systems," MIT Press, Cambridge, MA.

Holland, J. H. 1992. "Adaptation in Natural and Artificial Systems," Univ. of Michigan Press, Ann Arbor, MI.

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

Karlin, S., and McGregor, J. 1971. On mutation selection balance for two-locus haploid and diploid populations, *Theor. Popul. Biol.* **2**, 60–70.

Karlin, S., and Tavaré, S. 1981. The detection of a recessive visible gene in finite populations, *Genet. Res. Cambridge* **37**, 33–46.

Karlin, S., and Tavaré, S. 1982. A diffusion process with killing: The time to formation of recurrent deleterious mutant genes, *Stoch. Processes Appl.* **13**, 249–261.

Karlin, S., and Taylor, H. M. 1981. "A Second Course in Stochastic Processes," Academic Press, Boston, MA.

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

Michalakis, Y., and Slatkin, M. 1996. Interactions of selection and recombination in the fixation of negative-epistatic genes, *Genet. Res. Cambridge* **67**, 257–269.

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

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

Otto, S. P., Feldman, M. W., and Christiansen, F. B. 1994. Some advantages and disadvantages of recombination, *in* "Lecture Notes in Biomathematics" (S. A. Levin, Eds.), pp. 198–211, Springer-Verlag, Berlin/Heidelberg/New York.

Phillips, P. C. 1996. Waiting for a compensatory mutation: Phase zero of the shifting-balance process, *Genet. Res. Cambridge* **67**, 271–283.

Press, W. H., Saul, A., Teukolsky, W. T., Vetterling, and Flannery, B. P. 1992. "Numerical Recipes in C," 2nd ed. Cambridge Univ. Press, Cambridge, UK.

Wright, S. 1931. Evolution in Mendelian populations, *Genetics* **16**, 97–159.

Wright, S. 1977. "Evolution and the Genetics of Populations," Vol. 3, Univ. of Chicago Press, Chicago.

Zhivotovsky, L. A., Feldman, M. W., and Christiansen, F. B. 1994. Evolution of recombination among multiple selected loci: A generalized reduction principle, *Proc. Natl. Acad. Sci. USA* **91**, 1079–1083.