The Effect of Deleterious Alleles on Adaptation in Asexual Populations

Toby Johnson*,[1] and Nick H. Barton†

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

Manuscript received November 26, 2001
Accepted for publication May 17, 2002

ABSTRACT

We calculate the fixation probability of a beneficial allele that arises as the result of a unique mutation in an asexual population that is subject to recurrent deleterious mutation at rate U. Our analysis is an extension of previous works, which make a biologically restrictive assumption that selection against deleterious alleles is stronger than that on the beneficial allele of interest. We show that when selection against deleterious alleles is weak, beneficial alleles that confer a selective advantage that is small relative to U have greatly reduced probabilities of fixation. We discuss the consequences of this effect for the distribution of effects of alleles fixed during adaptation. We show that a selective sweep will increase the fixation probabilities of other beneficial mutations arising during some short interval afterward. We use the calculated fixation probabilities to estimate the expected rate of fitness improvement in an asexual population when beneficial alleles arise continually at some low rate proportional to U. We estimate the rate of mutation that is optimal in the sense that it maximizes this rate of fitness improvement. Again, this analysis relaxes the assumption made previously that selection against deleterious alleles is stronger than on beneficial alleles.

IT is often useful to view adaptive evolution in an asexual population (for example, on a nonrecombining chromosome) as two separate processes. The first process is the origin of new beneficial alleles by mutation, and the second process is the fixation of some of those alleles by natural selection. This article is concerned with the second process and specifically with calculating the probability of fixation of a beneficial allele, assuming that it starts at a low frequency. This problem was first studied by modeling the copy number of the beneficial allele as a branching process (Fisher 1922, 1930; Haldane 1927), assuming that the number of descendants from each copy of the beneficial allele is drawn from a Poisson distribution with mean W, where $W \equiv 1 + s_b$ is the absolute fitness of an individual carrying the beneficial allele. In a large population of fixed size N, the probability of ultimate fixation of a single copy of a beneficial allele is $P_{fix} = p[s_b]$ (for $1/N \ll s_b$), where $p[\cdot]$ is the unique function satisfying

$$p[s] \equiv 1 - \exp[-(1 + s)p[s]], \quad s \geq 0, p[s] \geq 0 \quad (1)$$

(Fisher 1922) and $P_{fix} \simeq 2s_b$ when $1/N \ll s_b \ll 1$ (Haldane 1927).

The geographical invariance principle states that the fixation probability of a beneficial allele is unaffected by spatial structuring of the population when there is no variation in W between demes (Maruyama 1970, 1971; Nagylaki 1982). However, almost any other departure from an idealized Wright-Fisher population at linkage equilibrium violates the assumptions made by Fisher (1922) and Haldane (1927). Specifically, their result does not hold when there is variation in W, and this is likely under a range of biologically realistic conditions. Examples include variation in space caused by local extinction and recolonization (Barton 1993; Barton and Whitlock 1997), variation in time caused by fluctuating population size (Ewens 1967; Otto and Whitlock 1997), variation across genetic backgrounds caused by selection at linked loci (Fisher 1930; Hill and Robertson 1966; Robertson 1970; Manning and Thompson 1984; Barton 1994, 1995; Charlesworth 1994; Peck 1994; Caballero and Santiago 1995; Gerrish and Lenski 1998; Stephan *et al.* 1999; Orr 2000b; Gerrish 2001; Rice and Chippindale 2001), and variation in the intensity of selection on the beneficial allele itself (for example, Pollak 1966b; Kimura and Ohta 1970; Barton 1987).

An accurate formula for fixation probabilities, based on a biologically appropriate model, is desirable for at least two reasons. First, it is one of the building blocks of more complex evolutionary models, in which the behavior of rare beneficial alleles is not explicitly modeled. Instead, the convenient assumption is made that a fraction P_{fix} of beneficial alleles reach frequencies large enough to actually be considered in the model, and the remaining fraction $(1 - P_{fix})$ are lost while still rare and can be ignored altogether (for example, Kimura 1983; Berg 1995; Hartl and Taubes 1996, 1998; Gerrish and Lenski 1998; Orr and Kim 1998; Orr 1998, 2000a,b;

[1] *Corresponding author:* Department of Zoology, University of British Columbia, 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada. E-mail: johnson@zoology.ubc.ca

Poon and Otto 2000; Gerrish 2001). In some models this assumption is in fact a limiting property of the model, obtained when selection is strong and mutation is weak relative to drift (Gillespie 1983a,b). Second, laboratory evolution experiments with *Escherichia coli* (Gerrish and Lenski 1998; Imhof and Schlötterer 2001) and vesicular stomatitis virus (Miralles *et al.* 1999) have allowed inferences to be made about the number of beneficial mutations that have arisen over the course of the experiment, on the basis of an observation of the number of selective sweeps that have occurred. Clearly, making accurate estimates requires an accurate formula for fixation probability (as pointed out by Orr 2000b).

This article is primarily concerned with how the fixation probability of a beneficial allele at one locus is influenced by segregating deleterious alleles at other loci, in a completely asexual population or along a completely nonrecombining chromosome. In the absence of recombination, beneficial mutations that arise in a given genetic background are effectively "trapped" in it, unable to recombine into other backgrounds (Fisher 1930, p. 122). Variation in fitness between the different backgrounds, caused by segregating deleterious alleles, can reduce or outweigh the advantage conferred by a beneficial allele, resulting in a substantially reduced fixation probability. The magnitude of this reduction has been the subject of several theoretical investigations (Manning and Thompson 1984; Charlesworth 1994; Peck 1994; Orr 2000b) and also of a recent empirical study (Rice and Chippindale 2001). This effect has also been studied theoretically for sexual populations, but with the emphasis on either a small number of tightly linked loci (Barton 1995; Stephan *et al.* 1999) or a large number of loosely linked loci (Peck 1994; Barton 1995).

Previous analytical work on asexual models has either assumed a fixed selection coefficient against deleterious alleles (s_d) with $s_b < s_d$ (Manning and Thompson 1984; Charlesworth 1994; Peck 1994) or has allowed s_d to vary across loci assuming that $s_b < \min\{s_d\}$ (Barton 1995; Orr 2000b). Under either of these assumptions, the frequency of the beneficial allele is expected to increase only if it is present in the most fit genetic background, as first pointed out by Fisher (1930, p. 122). If, prior to the origination of the beneficial allele, the frequency of the most fit background or genotype is f_0, and if a beneficial allele that originates on a lower fitness background has negligible probability of subsequently moving onto the most fit background, then

$$P_{\text{fix}} = f_0 p[s_b] \tag{2}$$

for $1/N \ll s_b < s_d$ (Charlesworth 1994; Peck 1994). When all deleterious alleles have a fixed effect, the value of f_0 in a population at equilibrium can be calculated by knowing that if the number of new deleterious alleles per individual per generation is Poisson distributed with mean U then the number of deleterious mutations per individual is Poisson distributed with mean U/s_d (Haigh 1978) and hence $f_0 = \exp[-U/s_d]$. Orr (2000b) proved that when selection coefficients against deleterious alleles are distributed with harmonic mean $E_h(s_d)$ then at equilibrium $f_0 = \exp[-U/E_h(s_d)]$. (This result was stated previously without proof by Gessler 1995 and Charlesworth 1996.)

In this article we assume deleterious alleles of fixed effect but relax the requirement that $s_b < s_d$. This case has been studied previously by Manning and Thompson (1984). However (as discussed by Peck 1994), their analysis made an erroneous assumption that $P_{\text{fix}} = 0$ if the beneficial allele was lost from the background in which it first arose, and also only a table of numerical results was presented. Peck (1994) conducted some simulation work for $s_b > s_d$. We present an algorithm for exact solution of the model studied by Manning and Thompson (1984) and Peck (1994). This solution is valid when $s_b \gg 1/N$ and $\exp[-U/s_d] \gg 1/N$. We also study the case where the beneficial mutation arises in a population that is not at equilibrium with respect to deleterious allele frequencies. A population will be perturbed from equilibrium by a selective sweep or a population bottleneck, and if selection against deleterious alleles is weak then the timescale of the approach back to mutation-selection equilibrium will be very slow (Johnson 1999). Indeed, in a rapidly evolving asexual population it is likely that frequencies of deleterious alleles will fluctuate continually and it is therefore of some interest to consider how fixation probabilities are altered in out-of-equilibrium populations. The more specific question addressed here is: What effect does the sweep to fixation of one beneficial allele have on the probability of subsequently arising beneficial alleles also becoming fixed?

We apply our results for fixation probabilities by estimating the expected rate of fitness improvement in an asexual population when beneficial alleles arise continually by mutation at rate kU per individual per generation, with $k \ll 1$. Our analysis is essentially an extension of the work of Orr (2000b), who used Equation 2 to estimate fixation probabilities and so derived results that are valid only when $s_b < s_d$. Following Orr (2000b), we estimate the rate of mutation that is optimal in the sense that it maximizes this rate of fitness improvement.

GENERAL METHODS FOR CALCULATING FIXATION PROBABILITIES

The branching process model as originally developed (Fisher 1922, 1930; Haldane 1927) assumes that W is the same for all copies of the beneficial allele and is constant over time. Building on work by Pollak (1966a, 1972), Barton (1995 and references therein) has applied the theory of multitype branching processes (see, *e.g.*, Harris 1963) to study how fixation probabilities

are influenced when W varies according to the "site" in which a given copy of the beneficial allele finds itself. This is a flexible approach in which sites can represent demes or microhabitats (Pollak 1966a, 1972; Barton 1987, 1993), genetic backgrounds (Barton 1994, 1995), age classes, or any other aspect of population structure. Relatively complex models can be analyzed in a straightforward way by considering the probability $Q_{i,t}$ that a single copy of the beneficial allele present in site i in generation t is ultimately *lost* from the population. An expression for $Q_{i,t}$ can be found by summing over all possible numbers of offspring and all possible movements of offspring between sites. If each copy of the allele present in site i at time t independently gives rise to n offspring, where n is drawn from a Poisson distribution with mean $W_{i,t}$, and the independent probability that each of these offspring will be in site j at time $t + 1$ is $m_{i,j}$, and each offspring in site j has independent probability $Q_{j,t+1}$ of being lost, then

$$Q_{i,t} = \sum_{n=0}^{\infty} e^{-W_{i,t}} \frac{(W_{i,t})^n}{n!} (Q_{i,t}^*)^n \equiv \exp[W_{i,t}(Q_{i,t}^* - 1)], \quad (3)$$

where

$$Q_{i,t}^* = \sum_j m_{i,j}(Q_{j,t+1}) \quad (4)$$

is the probability of loss given that the allele at time t has exactly one offspring (Barton 1995). These equations extend in a straightforward way to give the generating function for the distribution of the number of copies of the beneficial allele.

The assumption that separate copies of the allele are independent means that the branching process is an appropriate model only when the number of copies of the beneficial allele is small relative to the total population size. When $s_b \gg 1/N$ it is reasonable to assume that deterministic forces will prevail when the branching process model breaks down in this way, and in this case an allele that arises in background i at time t and that is never lost is said to become established. This occurs with probability $P_{i,t} = 1 - Q_{i,t}$. The probability of establishment for a beneficial mutation that arises in a random genetic background is denoted $P_{\text{fix},t}$ and is calculated as an average of the $P_{i,t}$, weighted by the probability of the beneficial mutation initially occurring in each background i. If an allele is established, it is not actually guaranteed fixation, but its frequency might approach a polymorphic equilibrium or it might become fixed only in some sites.

Because the branching process model assumes that the fate of an allele of interest is determined while it is rare, it cannot be used to calculate fixation probabilities for slightly deleterious mutations or for beneficial mutations that confer an advantage that is weak relative to the effects of genetic drift. For the same reason, branching process models cannot be used to determine the distribution of times taken until ultimate fixation of an allele. To address these types of questions Kimura (1957) used a diffusion approximation, which is valid when change is approximately continuous in time (large population sizes and weak selection). This approach has been used extensively to study the expected time to fixation (Kimura and Ohta 1969) and how fixation probabilities are influenced by population subdivision (Maruyama 1970, 1971; Nagylaki 1982; Barton 1987, 1993; Barton and Rouhani 1987), by "background selection" due to segregating deleterious alleles (Charlesworth 1994), or by changing population size (Otto and Whitlock 1997). Despite the advantages of the diffusion approximation, the branching process model remains a useful tool because it can often be analyzed much more easily.

MODEL

The model used is identical to the one studied by Manning and Thompson (1984) and Peck (1994). The notation used is summarized in Table 1. We consider the fate of a beneficial allele within a haploid population of fixed size N. Individuals without the beneficial allele are called "wild type." We assume that N is sufficiently large for deterministic results to apply to the wild-type subpopulation and for the effect of Muller's ratchet (loss of the most fit genotype by genetic drift; see Muller 1964; Haigh 1978; Gordo and Charlesworth 2000) within the wild-type subpopulation to be negligible over the timescale of interest. (A sufficient condition for this is $\exp[-U/s_d] \gg 1/N$.) We assume that deleterious mutations arise at a constant rate U per genome per generation, that there are an effectively infinite number of equivalent biallelic loci, and that at each locus the deleterious allele reduces fitness by a factor $(1 - s_d)$ with multiplicative effects across loci. Genetic backgrounds (sites) can therefore be enumerated by i, where i is the number of deleterious alleles and takes discrete values $i = 0, 1, 2, \ldots$. Back mutations are ignored. We first consider a wild-type population at mutation-selection equilibrium. We then consider the more complex scenario of a wild-type population that is initially free of deleterious alleles and is approaching equilibrium. The correspondingly more complex calculations have been placed in appendices a and b.

We assume that a single beneficial allele arises in a randomly chosen wild-type individual, and its presence increases relative fitness by a factor $(1 + s_b)$ regardless of the genetic background on which it is expressed; that is, we assume no epistasis for fitness. Except for its small size, the subpopulation carrying the beneficial allele is identical to the large wild-type (sub)population. That is, deleterious mutations arise at the same rate U and have the same effect on fitness $(1 - s_d)$. Because the number of copies of the beneficial allele is initially small, we consider the progress of Muller's ratchet within this subpopulation. We calculate the fixation probability for the beneficial allele, P_{fix}, by considering its copy number

TABLE 1

Frequently used notations

Symbol	Usage
C	Average improvement in log-fitness per generation.
ΔC	Average improvement in log-fitness per beneficial substitution.
f_i	Fraction of the wild-type population carrying i deleterious alleles.
h_i	Number of deleterious alleles fixed by hitchhiking, conditional on the beneficial allele arising in background i and becoming fixed.
i	Number of deleterious alleles in a given genetic background.
$i_{\max}$	Greatest i for which $P_i \neq 0$.
I	Inflation of fixation probability caused by nonequilibrium wild-type population.
k	Ratio of beneficial to deleterious mutations.
$m_{i,j}$	Probability that offspring is in site j given parent in site i.
N	Population size.
$p[\cdot]$	Unique function satisfying $0 \leq p[s] \equiv 1 - \exp[-(1 + s)p[s]]$.
p_i	Abbreviation for $p[e^{-U} W_i - 1]$.
P_{fix}	Fixation probability of a single copy of the beneficial allele arising in a random genetic background.
P_i	Fixation probability of a single copy of the beneficial allele in background i.
Q_i	Extinction probability; $1 - P_i$.
R	Relative fixation probability (relative to no interference).
s_{b}	Selection coefficient for the beneficial allele under consideration.
s_{d}	Selection coefficient against a deleterious allele.
U	Deleterious mutation rate per individual per generation.
w	Fitness relative to fittest wild-type genotype present.
W	Absolute fitness, *i.e.*, expected number of offspring.
x_i	Contribution to P_i arising from beneficial alleles lost from their original background but fixed in some lower fitness background.
λ	Mean i in the wild-type population.

in different genetic backgrounds as a multitype branching process. We assume a Poisson distribution of offspring number.

To estimate the long-term average rate of adaptation, measured as the rate of fitness improvement, we embed our model for fixation probabilities within a more complex model. This is a generalization of the model of Orr (2000b). We make the standard assumption that the beneficial mutation origination process is Poisson and occurs at rate kU per individual per generation, where $k \ll 1$ is the ratio of beneficial to deleterious mutations. Because the population size is N the total number of new beneficial alleles arising per generation is NkU. We then make the questionable assumption that each such mutation is fixed with equal and independent probability P_{fix}, so that the mutation fixation process is also Poisson and occurs at rate $NkUP_{\text{fix}}$. This assumption requires that the perturbations due to beneficial substitutions are small and transient enough that the fixation probabilities of most beneficial mutations are well approximated by the equilibrium population results. In making this assumption we also ignore interference between multiple beneficial alleles (the Hill-Robertson effect; Fisher 1930; Muller 1932; Hill and Robertson 1966; Felsenstein 1974; Barton 1995; Gerrish and Lenski 1998; Gerrish 2001). We assume that successively fixed beneficial alleles have multiplicative effects on fitness.

ANALYSIS

Fixation probabilities: Fitness relative to the fittest *wild-type* individual is denoted w. Hence, when the beneficial allele is present in the fittest possible individual it has relative fitness $w = w_0 = (1 + s_{\text{b}})$. A beneficial allele present in genetic background i has relative fitness

$$w_i = (1 + s_{\text{b}})(1 - s_{\text{d}})^i. \tag{5}$$

At this stage, a minor technical point should be made about two factors that have not been made very explicit in some previous analyses, although they were discussed in the Appendix of Peck (1994). In the branching-process model the expected number of offspring of a given genotype is its absolute fitness W_i, which with a constant population size is its fitness relative to the population mean fitness, $W_i = w_i/\overline{w}$. Because the beneficial allele is rare by assumption, $\overline{w}$ is equal to the mean fitness of the wild-type subpopulation, and because we assume here that the wild-type population is at equilibrium $\overline{w} = e^{-U}$ (Kimura and Maruyama 1966) and therefore

$$W_i = w_i e^U. \tag{6}$$

The importance of this difference between absolute and relative fitness is not necessarily apparent when both the wild-type population is at equilibrium *and* only un-

mutated offspring are of interest. A fraction e^{-U} of the offspring of a given individual carrying the beneficial allele are free from additional deleterious mutations, and so in this special case these two factors cancel out and correct results can be derived by assuming that the "effective absolute fitness" is $e^{-U}w_i e^{U} = w_i$. This simplification does not apply generally, however.

In Equations 3 and 4, $Q_{i,t}$ is the probability of loss of a single copy of the beneficial allele present after selection followed by movement between sites. Since here "movement between sites" represents deleterious mutation, the probability of the beneficial allele arising in site i is calculated using the frequencies of the different sites after deleterious mutation. Because the number of deleterious alleles, i, carried by a randomly chosen wild-type individual is Poisson distributed with mean $\lambda = U/s_d$ (Haigh 1978), we have that

$$f_i = \frac{\lambda_t^i e^{-\lambda}}{i!} \tag{7}$$

is the probability of the beneficial mutation arising with i deleterious alleles (*i.e.*, the frequency of background i). Equation 7 holds only when $\exp[-U/s_d] \gg 1/N$.

It is more convenient to rewrite Equation 4 in terms of the probability P_i^* that an allele in background i is never lost (where $P_i^* = 1 - Q_i^*$). The probability that a given copy of the beneficial allele in site i is moved to site $i + j$ by deleterious mutation is simply

$$m_{i,i+j} = \frac{e^{-U}U^j}{j!}. \tag{8}$$

When we substitute (4) and then (8) into (3) we obtain

$$P_{i,t} = 1 - \exp\left[-W_{i,t}\sum_{j=0}^{\infty}\frac{e^{-U}U^j}{j!}P_{i+j,t+1}\right]. \tag{9}$$

As mentioned briefly earlier in this article and discussed at more length by Peck (1994), there may be backgrounds in which the beneficial allele confers a net advantage in the short term but that will ultimately give rise to a lineage with a lower mean fitness than the wild-type population; such beneficial alleles are assumed not to ultimately fix. Consider a beneficial allele subpopulation that has become large and that includes individuals with i, $i + 1$, $i + 2$, . . . deleterious alleles. The most fit individual in this subpopulation has fitness $\max\{w_B\} = (1 + s_b)(1 - s_d)^i$ and so the subpopulation will ultimately approach mutation-selection balance with $\overline{w}_B = (1 + s_b)(1 - s_d)^i e^{-U}$. The wild-type subpopulation ultimately approaches $\overline{w}_W = e^{-U}$ and if $\overline{w}_B > \overline{w}_W$, equivalently $(1 + s_b)(1 - s_d)^i > 1$ or

$$i \le i_{\max} = \left\lfloor \frac{\ln[1 + s_b]}{-\ln[1 - s_d]} \right\rfloor, \tag{10}$$

then the beneficial allele subpopulation will ultimately replace the wild-type subpopulation. In condition (10) $\lfloor\cdot\rfloor$ denotes the integer part and $i_{\max}$ is the largest value of i where the condition is satisfied, noting that it is always satisfied for $i = 0$. Here we have assumed that the wild-type population is sufficiently large that the beneficial allele subpopulation does not have time to fix before reaching approximate mutation-selection balance, and we have ignored back mutation of deleterious alleles. (The validity of these assumptions is discussed below.) To find the probability that the beneficial allele ultimately fixes we need to follow only the number of copies in backgrounds $0 \le i \le i_{\max}$, because if it is lost from all of these backgrounds then it can never ultimately fix. Therefore we can replace the ∞ in the upper limit of the sum in Equation 9 with $(i_{\max} - i)$.

As $t \to \infty$ with N constant the state of the wild-type population becomes constant over time, and at equilibrium both $P_{i,t+1}$ and $P_{i,t}$ converge to the single value P_i (an abbreviation for $P_{i,\infty}$).

By making the simplifications described in the previous two paragraphs, we obtain a set of simultaneous equations in P_i for $i = 0, 1, 2, \ldots, i_{\max}$. In general, we can start by solving

$$\begin{aligned} P_{i_{\max}} &= 1 - \exp[-e^{-U}W_{i_{\max}}P_{i_{\max}}] \\ &= 1 - \exp[-w_{i_{\max}}P_{i_{\max}}] \\ &= p_{i_{\max}}, \end{aligned} \tag{11}$$

where p_i is an abbreviation for $p[w_i - 1]$ and $p[s]$ is the fixation probability of a Poisson branching process with mean $1 + s$ (see Equation 1). Each P_i can then be calculated numerically in descending order of i by solving

$$P_i = 1 - \exp\left[-w_i\sum_{j=0}^{i_{\max}-i}\frac{U^j}{j!}P_{i+j}\right]. \tag{12}$$

Once the P_i are known, the net fixation probability P_{fix} can be calculated by averaging over all of the different genetic backgrounds

$$P_{\text{fix}} = \sum_{i=0}^{i_{\max}} f_i P_i, \tag{13}$$

where the f_i are given by setting $\lambda = U/s_d$ in Equation 7. There does not appear to be a more concise general expression for P_{fix}, except for the special case where $i_{\max} = 0$, which was solved by Peck (1994; see Equation 2 above). The calculations described here have been automated in a *Mathematica* (Wolfram 1996) notebook available from T. Johnson on request.

The rate of adaptation: Our model for long-term adaptation assumes that beneficial mutations are rare enough that they can be considered, to a good approximation, to arise in close-to-equilibrium populations, so that P_{fix} is relevant. We wish to find an approximation for C, the expected increase in mean log-fitness per generation. To do this we first find an approximation for ΔC, the expected increase in log-fitness per beneficial substitution, which in turn requires an approximation for ΔC_i, the expected increase in log-fitness per benefi-

cial substitution conditional on the beneficial allele arising in background i.

If the beneficial allele arises on background i and is ultimately fixed, then some number h_i of deleterious alleles will be fixed by hitchhiking (allowing the possibility of $h_0 = 0$). Clearly $h_i \geq i$, but the problem is that h_i conditional on fixation is a random variable and may exceed i in the event that the beneficial mutation is lost from the background on which it originally arose but ultimately does becomes fixed. We make use of the decomposition $P_i = p_i + x_i$, which is detailed in APPENDIX C. Here $p_i = p[w_i - 1]$ is the probability of fixation given that the beneficial allele fixes in the background in which it arose (i), and x_i is the probability of fixation given that the beneficial allele is lost from the background in which it arose. x_i can be calculated directly by numerical solution of Equation C2 or simply by taking the difference between P_i and p_i. This partitioning of fixation events into two mutually exclusive possibilities suggests that for $U \ll 1$ it might be reasonable to suppose that the fate of a beneficial allele lost by mutation from background i, conditional on eventual fixation, is identical to the fate of a beneficial allele that arises in background $i + 1$, again conditional on ultimate fixation. This leads to the approximation

$$E(h_i) \simeq \frac{1}{P_i}(p_i i + x_i E(h_{i+1})), \tag{14}$$

which can be calculated in decreasing order of i because $x_{i_{\max}} = 0$. The expected increase in population mean fitness given that a beneficial allele arising on background i becomes fixed can be similarly approximated, and because for small s_b the mean increases in fitness and log-fitness are roughly equal we obtain

$$\Delta C_i \simeq \frac{1}{P_i}(p_i((1 + s_b)(1 - s_d)^i - 1) + x_i \Delta C_{i+1}). \tag{15}$$

Both of these quantities can be averaged over the distribution of $f_i P_i$. Then we obtain the expected number of deleterious alleles that will hitchhike with a beneficial allele arising on a random background, conditional on its ultimate fixation

$$E(h) \simeq \frac{1}{P_{\text{fix}}} \sum_{i=0}^{i_{\max}} f_i P_i E(h_i) \tag{16}$$

and the expected increase in mean log-fitness per fixation

$$\Delta C \simeq \frac{1}{P_{\text{fix}}} \sum_{i=0}^{i_{\max}} f_i P_i \Delta C_i. \tag{17}$$

As explained above, the assumption of our model for continued adaptation is that the mutation fixation process is Poisson and occurs at rate $NkUP_{\text{fix}}$. Because we assume multiplicative effects of multiple beneficial alleles, the long-term average rate of mean log-fitness increase is given approximately by

$$C \simeq NkUP_{\text{fix}} \Delta C. \tag{18}$$

When U is varied and the other model parameters are held constant, C is maximized at a particular value of U, which we call the optimum mutation rate U_{opt}. ORR (2000b) proved that when $s_b < s_d$ (so that $i_{\max} = 0$) the optimum mutation rate is given simply by $U_{\text{opt}} = s_d$. This result can be seen by substituting (2) into (18) and differentiating to obtain

$$\frac{dC}{dU} \simeq Nke^{-U/s_d}\left(1 - \frac{U}{s_d}\right) 2 s_b \Delta C \tag{19}$$

(because $\Delta C \simeq s_b$ does not depend on U for $i_{\max} = 0$) and noting that $dC/dU = 0$ and $d^2C/dU^2 < 0$ when $U = s_d$ (ORR 2000b). For $i_{\max} > 0$ no such simple derivation is possible, but U_{opt} can be found by numerical calculation reasonably efficiently using a golden section search, automated in a *Mathematica* (WOLFRAM 1996) notebook available from T. Johnson on request. It is notable that because both P_{fix} and ΔC are only functions of U, s_b, and s_d and do not depend on N or k, then U_{opt} must be a function of s_b and s_d only and also will not depend on N or k.

NUMERICAL RESULTS

Fixation probabilities when many genetic backgrounds are relevant: As we described in the Introduction (see also FISHER 1930; MANNING and THOMPSON 1984; CHARLESWORTH 1994; PECK 1994; BARTON 1995; ORR 2000b), when $s_b \leq s_d$ a beneficial allele must arise in the single most fit genetic background to have any chance of fixation, and the net fixation probability is given by Equation 2. However, when $s_b \gg s_d$ there are many genetic backgrounds in which a beneficial allele can arise and have some probability of fixation and so the situation is more complex. This is illustrated in Figure 1, which shows P_i (top, line), the fixation probability of a beneficial allele arising in a background with i deleterious alleles, assuming a wild-type population at equilibrium and parameter values $s_b = 5 \times 10^{-3}$, $s_d = 5 \times 10^{-4}$, and $U = 3 \times 10^{-3}$. The probability of arising in each background, f_i (top, dots), and the way that P_i and f_i combine (bottom) to determine the net fixation probability P_{fix} are also shown. For beneficial alleles arising in the least fit background of relevance, with $i = i_{\max} = 9$ deleterious alleles, the fixation probability is very small ($w_{i_{\max}} < \sim 1 + s_d$ and hence $P_{i_{\max}} < \sim 2s_d$ for $s_d \ll 1$) because the advantage of the beneficial allele is almost totally eliminated by the deleterious alleles it is linked to and because any new deleterious mutation will eliminate that advantage altogether. As i decreases P_i increases and approaches an asymptote, which is $P_0 \simeq 2(U + s_b) = 1.6 \times 10^{-2}$ for $U \ll 1$, $s_b \ll 1$. This asymptote occurs because the beneficial allele is in a genotype with absolute fitness approaching $(1 + s_b)\exp[U]$ and because many new deleterious mutations must occur to reduce

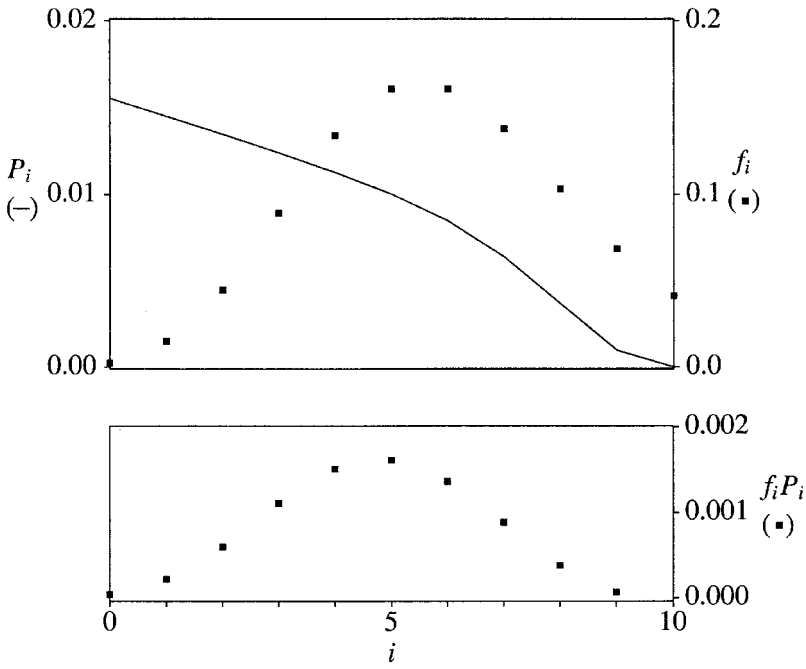

FIGURE 1.—A beneficial allele has probability f_i (top, dots) of arising in a background of i deleterious alleles. Given that it does, it has fixation probability P_i (top, line). The sum of f_iP_i (bottom) over all backgrounds gives the net fixation probability $P_{\text{fix}} \simeq 7.8 \times 10^{-3}$ for the parameter values used here ($s_b = 5 \times 10^{-3}$, $s_d = 5 \times 10^{-4}$, $U = 3 \times 10^{-3}$).

that advantage. A mathematical description of this behavior, which is a good approximation when $U \gg s_b$, is detailed in APPENDIX C. In this example the main contribution to the net fixation probability P_{fix} is from moderate fitness backgrounds with i intermediate between zero and i_{max}. For the parameter values used in Figure 1 a beneficial allele that fixes is most likely to be one that arose on a background of $i = 5$ deleterious alleles, and therefore at least one-half its selective advantage will be negated by the deleterious alleles that hitchhike to fixation with it.

Fixation probabilities in an equilibrium population: Figure 2 shows results for the situation where the beneficial allele arises in a population at equilibrium under selection and deleterious mutations of fixed effect. It explores the region of the parameter space where $U/s_d \leq 10$, so that $f_0 \geq e^{-10} \simeq 4.5 \times 10^{-5}$ will represent a large number of individuals for population sizes that are realistic, at least for bacteria. Shown is the relative fixation probability when there is interference, $R = P_{\text{fix}}/p[s_b]$, where the fixation probability of the beneficial allele is P_{fix} and the value it would take in the absence of any interference is $p[s_b]$. A built-in *Mathematica* algorithm (WOLFRAM 1996) was used to fit contours to a lattice of values calculated by numerical solution of Equations 11 and 12. The irregularities in the contour lines are not caused by numerical inaccuracies but occur near where the number of relevant genetic backgrounds i_{max} changes from one integer value to the next, which causes the gradient of P_{fix} to change abruptly.

For the region of the parameter space considered in Figure 2 the relative fixation probability spans almost five orders of magnitude. A logarithmic scale of fixation probabilities is appropriate if one wishes to understand in which regions of the parameter space evolution essentially cannot proceed. However, for other questions a linear scale is more appropriate. In the case of competing subpopulations, a small difference in rates of adaptation is critical, and the rate of adaptation is approximately linear in the fixation probability. In a log-log-linear space the (s_b, s_d, R) surfaces show regions with roughly constant either high ($R \simeq 1$) or low ($R \simeq 0$) fixation probability. To help visualize this we show additional contour lines at $R = 0.9$ (dotted lines) and $R = 0.5$ (dashed lines), which together with the first solid contour at $R = 0.1$ roughly define the transition from no effect to a severe effect of interference from deleterious alleles.

In all of Figure 2, the area above and to the left of the diagonal $s_b = s_d$ represents the region of the parameter space where Equation 2 applies, and the area below and to the right of this diagonal represents the region of the parameter space where the methods developed above are necessary to calculate the fixation probability.

Several trends worthy of comment are evident in Figure 2. First, by comparing across all three panels we can see that fixation probabilities are reduced more severely by interference as the rate of deleterious mutation U increases. This result is expected. Second, the severity of the reduction in fixation probability often *but not always* increases monotonically as the strength of selection against deleterious alleles s_d decreases. An example of nonmonotonicity in s_d can be seen when $U = 10^{-3}$ and $s_b = 10^{-2.5}$ (Figure 2, top). This perhaps counterintuitive result occurs because there are two opposing forces at work here. As s_d decreases the frequency of deleterious alleles increases, and hence they are more likely to be present in the background on which the beneficial allele

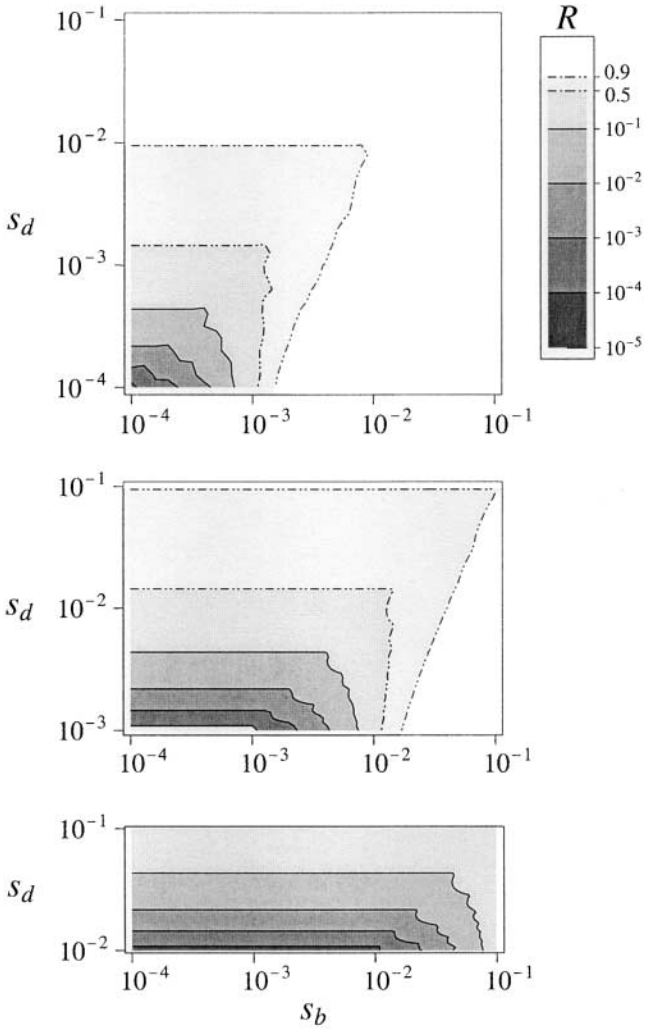

FIGURE 2.—The fixation probability P_{fix} of a beneficial allele conferring a selective benefit s_b is reduced by interference from segregating deleterious alleles causing disadvantage s_d that arise at completely linked loci at rate U. The relative fixation probability in an equilibrium population ($R = P_{fix}/p[s_b]$, measured relative to the fixation probability without interference) is shown as a function of s_b and s_d. From top to bottom $U = 10^{-3}$, $U = 10^{-2}$, and $U = 10^{-1}$.

arises, but the effect of any one deleterious allele in reducing the advantage of the beneficial allele is less. The graphs show that, for the parameter space explored, the former force tends to dominate the latter. For $s_b < s_d$ examination of Equation 2 shows that R increases as U increases and s_d decreases, but it was not at all clear that this dependency would be true for most but not all s_b.

A third visible trend is that beneficial alleles of larger effect have fixation probabilities that are less influenced by segregating deleterious alleles. This is an intuitively reasonable result, but one that is *not* true when $s_b < s_d$. Equation 2 shows that R is independent of s_b, which can be seen to be true in the parts of Figure 2 where this equation applies, *i.e.*, everything above and left of the line $s_b = s_d$. To the right of and below the line we see that the situation is more complex. The final trend worth noting is that in some parts of the parameter space there is a catastrophic reduction in fixation probability caused by interference from segregating deleterious alleles. To a coarse approximation, it can be said that $R \simeq 0.5$ when $s_b = \max\{s_d, U\}$ and R will be very small when both $s_b \ll s_d$ and $s_b \ll U$.

The three parts of Figure 2 are almost perfect replicas of each other, offset by the value of U, suggesting that R depends only on two compound parameters s_b/U and s_d/U. Equation 2 shows that this is true when $s_b < s_d$. In APPENDIX C we show that this is also true when $s_d \ll s_b \ll 1$ and $U \ll 1$, that is, when selection and mutation are both weak and many genetic backgrounds are relevant.

Nonequilibrium populations: In APPENDICES A and B we derive results for fixation probabilities when the population of interest was free of segregating deleterious alleles at some time $t = 0$ in the past. This initial condition approximates the effect of a rapid selective sweep, a severe population bottleneck with rapid recovery, or the founding of a laboratory evolution experiment from a single clone. The fixation probability of a beneficial mutation that arises at some subsequent time $t = \tau$ is denoted $P_{fix,\tau}$. The results above for an equilibrium population are a special limiting case of this scenario, where $\tau \to \infty$.

Consider first the simplest case $\tau = 0$. At that time a beneficial mutation is guaranteed to arise in a background free of deleterious alleles ($f_0 = 1$), which would increase its net fixation probability. However, at the same time the population mean fitness $\overline{w}$ would be unity, causing the absolute fitness of genotypes containing the beneficial allele to be lower than in an equilibrium population (where $\overline{w} = e^{-U}$; see Equation 6), which would cause a decrease in its net fixation probability. There are therefore two opposing forces at work, and their combined effect on the fixation probability of a beneficial allele is not obvious. It is not necessarily sufficient to argue that the net fixation probability is reduced by variance in fitness across backgrounds and is therefore higher when there is zero variance at $\tau = 0$, because this argument does not take into account the changing mean fitness of the wild-type population. (Indeed a variance-based argument fails to predict how P_{fix} depends on s_d in an equilibrium population.)

In APPENDIX B we prove that, for the special case where $s_b < s_d$, the fixation probability for a beneficial mutation occurring at time $\tau < \infty$ is always greater than in an equilibrium population ($\tau = \infty$). To see the importance of this result, consider further the case of a weakly selected beneficial allele with $s_b \ll U$, $s_b \ll s_d$ arising at time $\tau = 0$. Genotypes containing such a beneficial allele all give rise to, on average, less than one offspring of the same genotype at $t = 0$ and for some period of time thereafter (because $e^{-U}W_{i,0} < 1$, see APPENDIX A). To have any probability of fixation at least one copy of the

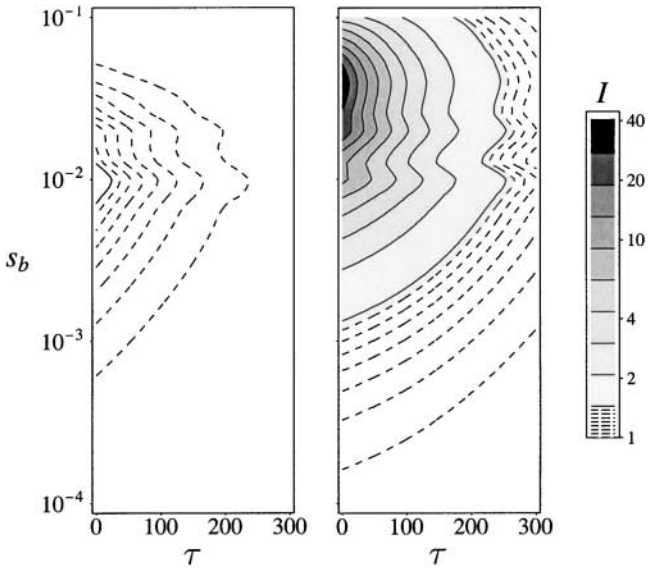

FIGURE 3.—The fixation probability P_{fix} of a beneficial allele conferring a selective benefit s_b is inflated if it arises some short time τ after segregating deleterious alleles have been purged from an asexual population. The degree of inflation ($I = P_{\text{fix},\tau}/P_{\text{fix},\infty}$, measured relative to the fixation probability in a population at mutation-selection equilibrium) is shown as a function of τ and s_b. Left, $U = 10^{-2}$, $s_d = 10^{-2}$, and the maximum inflation is $I \simeq 1.58$. Right, $U = 10^{-1}$, and $s_d = 10^{-2}$, and the maximum inflation is $I \simeq 37.7$. In both sides the same shading scheme is used, solid contours are at geometric intervals of $\sim$1.44, and dashed contours are at geometric intervals of $\sim$1.048.

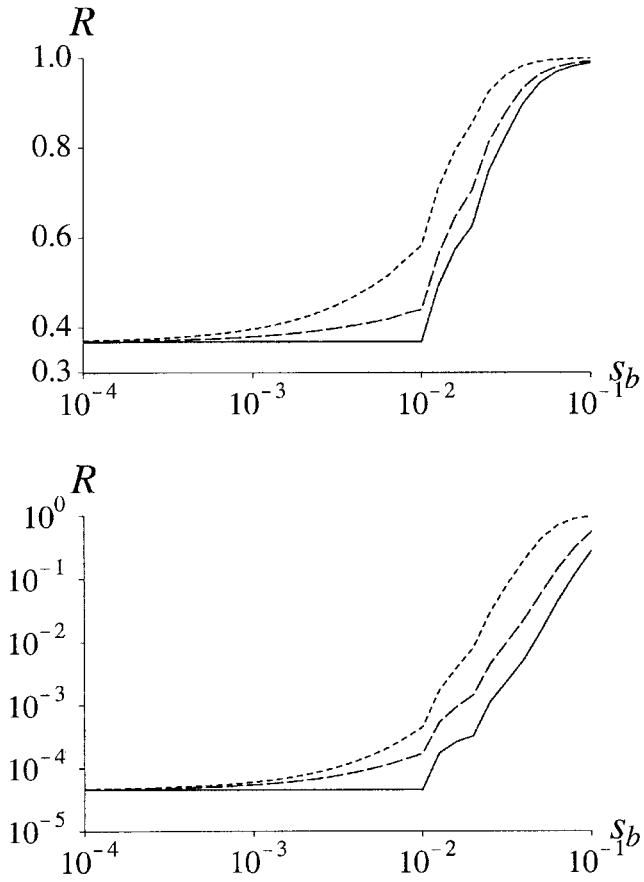

FIGURE 4.—Relative fixation probabilities, R, in a nonequilibrium population. Top, $s_d = 10^{-2}$, $U = 10^{-2}$; bottom, $s_d = 10^{-2}$, $U = 10^{-1}$. Note the very different scales of the y-axes. Curves show $\tau = 0$ (dotted), $\tau = 100$ (dashed), and $\tau = 1000 \simeq \infty$ (solid).

allele must persist until the wild-type mean fitness has decayed significantly to have absolute fitness greater than one. It is therefore quite surprising to find that such beneficial alleles always have greater fixation probability than they would if they arose in an equilibrium population. We offer the following explanation for our perhaps counterintuitive result. As we go backward in time there is a decrease in fixation probability due to increasing $\overline{w}_t$ and an increase in fixation probability due to increasing $f_{0,t}$. Because $f_{0,t+1}/f_{0,t} = e^{-U}/\overline{w}_t$ these two forces are coupled, and the nature of this coupling ensures that, working backward in time from an equilibrium fixation probability, the fixation probability will always increase. On the basis of our numerical results for $s_b > s_d$ we speculate that this is true for all values of s_b and s_d.

We measure the effect of a nonequilibrium population by the inflation in fixation probability, $I = P_{\text{fix},\tau}/P_{\text{fix},\infty}$, measured relative to an equilibrium population. Figure 3 shows I as a function of s_b and of τ. In these calculations we assumed that after 1000 generations the population would be close to equilibrium. The choice of $s_d = 10^{-2}$ for these plots was influenced by the computer time required for the calculations (see APPENDIX A), and the choices of $U = 10^{-2}$ (left plot) and $U = 10^{-1}$ (right plot) represent situations where there is a moderate and a large reduction in fixation probability at equilibrium. For these parameter values, any inflation in fixation probability is a relatively short-lived effect and is negligible [in the sense that $(I - 1)/(\max\{I\} - 1) < 0.05$] after 300 generations. This is to be expected because a population perturbed from mutation-selection balance decays toward its equilibrium state on a timescale proportional to $1/s_d = 100$ generations (JOHNSON 1999). What is more remarkable is that the inflation in fixation probability occurs only for beneficial alleles with certain selection coefficients. For $U = 10^{-2}$ only beneficial alleles with $s_b \simeq 10^{-2}$ have inflated fixation probabilities and the effect is moderate ($I \leq 1.6$). For $U = 10^{-1}$ beneficial alleles with selection coefficients $s_b \simeq 10^{-2}$ or greater have inflated fixation probabilities and the effect is substantial ($I \simeq 40$). To understand why this is so, it is necessary to consider these fixation probabilities relative to the case with no interference, measured by $R = P_{\text{fix},\tau}/p[s_b]$, which are shown in Figure 4. The abrupt changes in gradient visible in the graph are caused when the number of relevant genetic backgrounds $i_{\max}$ changes from one integer value to the next.

Assuming that the fixation probability is always less

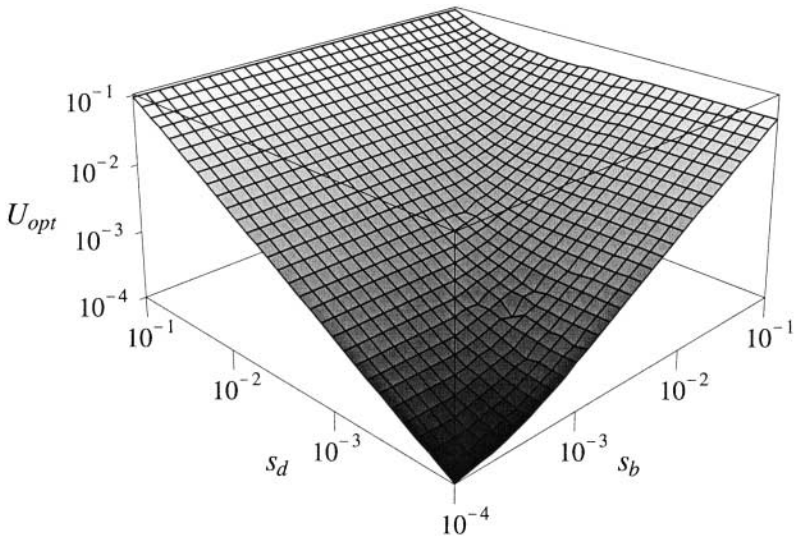

FIGURE 5.—Under the model assumptions (see text and Figure 6), the optimum deleterious mutation rate U_{opt} maximizes the long-term average rate of fitness improvement in an asexual population. U_{opt} depends only on s_b and s_d, against which it is plotted here.

than it would be in the complete absence of deleterious mutation (we prove this for $s_b < s_d$ in APPENDIX B and speculate that it is true always), then if fixation probabilities are only moderately reduced, as in the case $s_d = 10^{-2}$, $U = 10^{-2}$ (Figure 4, top), then the inflation in fixation probability I can be at most moderate. On the other hand when fixation probabilities are substantially reduced, as in the case $s_d = 10^{-2}$, $U = 10^{-1}$ (Figure 4, bottom), then the inflation in fixation probability I can also be substantial. The slightly mysterious peak of inflation I at $s_b \simeq 10^{-2}$ for $U = 10^{-2}$ can be partly explained because, for $s_b \gg 10^{-2}$, there is no reduction in fixation probability at equilibrium and hence there can be no transient inflation.

Figures 3 and 4 show that, for the specific departure from equilibrium that we have studied, there is little effect on fixation probability for weakly selected beneficial alleles. It is possible that this is because the fate of such alleles takes a long time to be determined, and the transient nonequilibrium state of the wild-type population is therefore of little relevance.

The rate of adaptation and the optimum mutation rate: Figure 5 shows numerical calculations of the optimum mutation rate U_{opt} as a function of the two parameters on which it depends, s_b and s_d. The optimum mutation rate is the rate that maximizes the long-term average rate of fitness increase, as estimated using Equation 18. These results are in agreement with ORR's (2000b) finding that $U_{opt} = s_d$ when $s_b \leq s_d$ (Figure 5, left). However, when $s_b > s_d$ the dependence on s_d is much weaker, and to a very coarse approximation $U_{opt} \simeq \max\{s_b, s_d\}$ for the whole parameter space examined here. This result can be explained in terms of our results for fixation probabilities. For any chosen combination of s_b and s_d imagine the appropriate points in Figure 2. Consider first a very small value of U. The fixation probability is high. Now imagine increasing U so that a "hole" starts to appear in the corner of the plane. At first, there is a less-than-linear decline in fixation probability with U and the rate of beneficial mutations increases linearly with U, so the rate of fitness improvement C increases. Suddenly, when U exceeds $\max\{s_b, s_d\}$ there is a catastrophic decline in fixation probability and an associated decline in C. Hence the rate of adaptation is maximized just before this catastrophe, at $U_{opt} \simeq \max\{s_b, s_d\}$. This is illustrated in Figure 6.

DISCUSSION

We have described a method for calculating fixation probabilities when segregating deleterious alleles of fixed effect jointly influence the fate of a beneficial allele in an asexual population. Our analysis includes as a special case the situation studied previously where any single deleterious allele overwhelms the advantage of the beneficial allele (MANNING and THOMPSON 1984; CHARLESWORTH 1994; PECK 1994; ORR 2000b). However, our analysis applies for any strength of selection on both the deleterious alleles and the beneficial allele, provided that selection is strong relative to drift and that Muller's ratchet is not operating rapidly. In particular, it allows study of the fate of a beneficial allele of large effect, where "large" means larger than the effect of a single deleterious allele. Our method allows fixation probabilities to be calculated to arbitrary precision by solution of a set of equations, but we cannot in general write down an equation for the fixation probability that improves our understanding of the effect we are studying. The value of our numerical results is that a reasonable intuition about the effect of interfering deleterious alleles on fixation probabilities can be developed by studying Figure 2. This numerical work together with the approximate analytical results derived in APPENDIX C justifies the following "rules of thumb" concerning $R = P_{fix}/p[s_b]$, a statistic that describes the fixation probability in the presence of segregating deleterious alleles

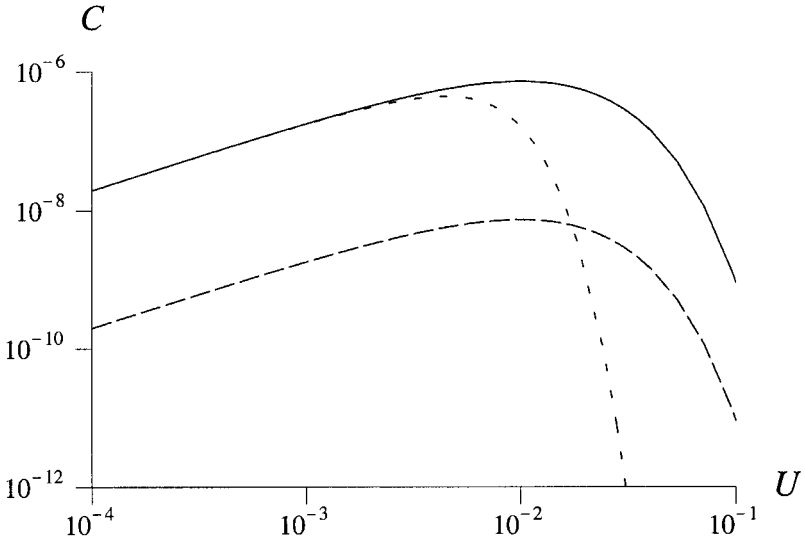

FIGURE 6.—Assuming the ratio of beneficial to deleterious mutations, k, is fixed and small, the long-term average rate of fitness improvement C in an asexual population is maximized when the deleterious mutation rate U takes a particular value, U_{opt}, which we term the optimum. U_{opt} is independent of k and of the population size N. The optimum is clear on plots of C against U, shown here for three choices of parameter values: $s_b = s_d = 10^{-2}$ (solid); $s_b = 10^{-2}$, $s_d = 10^{-3}$ (dotted); and $s_b = 10^{-3}$, $s_d = 10^{-2}$ (dashed). For all curves $Nk = 1$.

(P_{fix}) in an equilibrium population, relative to what it would be in the absence of deleterious alleles ($p[s_b]$).

1. For weak selection and mutation R depends only on the compound parameters s_b/U and s_d/U, which describe the strength of selection relative to the rate of deleterious mutation.
2. The relative fixation probability R cannot be predicted from only the variance in fitness in the wild-type population, which is Us_d.
3. For $s_b < s_d$ the relative fixation probability is $R = \exp[-U/s_d]$.
4. For $s_d \ll U$ and $s_b > U$ there is negligible reduction in fixation probability and $R \simeq 1$.
5. For $s_d \ll U$ and $s_b < U$ there is substantial reduction in fixation probability and $R \simeq 0$.

Conclusions 1, 2, 4, and 5 are novel.

Previous studies of the effect of linked deleterious alleles on fixation probabilities have all assumed $s_b < \min\{s_d\}$, with three relatively minor exceptions. PECK (1994) presented simulation results where $s_b > s_d$, but only for two choices of parameter values and thus gave little indication of how P_{fix} depends on model parameters. The earlier numerical work of MANNING and THOMPSON (1984) was similarly cursory, and their analysis contained an error (as discussed by PECK 1994 and described above also). ORR (2000b, Figure 2) showed simulation results where both s_b and s_d were drawn from continuous distributions for each mutational event and in a fraction ($\sim$2.5%) of pairwise cases $s_b > s_d$.

The emphasis on $s_b < s_d$ in previous studies has been influenced by a combination of mathematical convenience and the belief that a major component of adaptation is due to beneficial alleles of small effect. The view that s_b is typically small was argued by FISHER (1930) on the basis of a "geometrical" model, which involves evolution toward an optimum in high-dimensional phenotype space. In this model, mutations are assumed to alter the phenotype in a random "direction" and the probability that a given mutation produces a phenotype closer to the optimum increases as the effect of that mutation becomes smaller. Thus, the density function for the distribution of s_b (conditional on $s_b > 0$) is expected to be monotonically decreasing. This tendency becomes more marked as the dimensionality of phenotype space increases. Later studies (KIMURA 1983, pp. 154–155; ORR 1998, 1999, 2000a) describe the distribution of s_b conditional on fixation and so do not alter FISHER's (1930) conclusion concerning the distribution of effects of newly arising beneficial alleles. We note that these later studies have all assumed $P_{fix} \simeq 2s_b$, and so a possible area for further work is to examine expressions for fixation probability that are nonlinear in s_b, such as those derived here, in the context of Fisher's geometrical model.

There is an accumulating body of empirical evidence (reviewed by ORR 1999) suggesting that adaptation involves factors of moderate or large effect and so a complete theory of the population genetics of adaptation must include the parameter space where $s_b > s_d$. Whether more general evolutionary theories that deal with the evolution of sex (*e.g.*, CHARLESWORTH 1994; PECK 1994) or the degeneration of Y chromosomes (*e.g.*, CHARLESWORTH 1996; ORR and KIM 1998) must also deal with this part of the parameter space is difficult to answer. Clearly the absolute magnitude of s_b is of little relevance, and what matters (for models with fixed s_d) is whether $s_b < s_d$ is typical or at least provides a good approximation. It can be argued that $s_b < s_d$ for most mutations, because there is a class of deleterious mutations with $s_d \simeq 0.02$ and a class of lethal mutations (see, *e.g.*, MUKAI 1964; KEIGHTLEY 1996; KEIGHTLEY and EYRE-WALKER 1999; LYNCH *et al.* 1999). However, deleterious alleles of large or lethal effect segregate at low frequency or not at all and hence have little influence on the fixation probability of beneficial alleles. It may be more impor-

tant that there is a substantial "nearly neutral" class of deleterious mutations with very small effects on fitness (Ohta 1973, 1992; Ohta and Gillespie 1996; Davies *et al.* 1999; Keightley and Eyre-Walker 1999). It seems certain that there will be beneficial alleles that have greater selective advantage than any individual deleterious mutation in this latter class, but the combined effects of many such very slightly deleterious mutations could still have a substantial effect on fixation probabilities. The analysis for large s_b presented here may therefore be of importance for beneficial alleles that would not be considered (from a Fisherian viewpoint) to have a particularly large effect. However, whether $s_b < s_d$ or $s_b > s_d$ for "most" mutations (or pairs of mutations) may not be the relevant question. In reality s_d varies across loci and the effects of newly arising deleterious alleles are drawn from an approximately continuous distribution. Fixation probabilities for beneficial alleles depend on the approximately continuous distribution of fitnesses of genetic backgrounds on which beneficial alleles arise and on the way that subsequent deleterious mutation moves beneficial alleles from one background to another after they arise. An analysis of such a model will be described in a subsequent article.

In the results we remarked that there is a region of the parameter space where fixation probabilities are catastrophically reduced. For weak selection against deleterious alleles, $s_d \ll U$, there is a transition between high and low fixation probabilities that occurs in the region around $s_b \simeq U$. This transition becomes more marked as s_d becomes smaller and in fact approaches a step function in the limit $s_d \to 0$. This conclusion and in fact any application of our model when $s_d \ll U$ must be carefully qualified. As s_d becomes smaller the distribution of fitness in the wild-type population piles up around $w = e^{-U}$ with very small variance. In our analysis, we measure the fitness of the beneficial allele relative to the fitness of an individual with $w = 1$, and such individuals become vanishingly rare as $s_d \to 0$ with constant U. Technically speaking this would not be a problem if we held by the stated assumption that the distribution of fitness in the wild-type population is constant over time, but clearly the validity of this assumption becomes questionable when $f_0 = \exp[-U/s_d] \gg 1/N$. A second reason for our argument breaking down as $s_d \to 0$ is that we assume the beneficial allele subpopulation reaches mutation-selection equilibrium before displacing the wild-type subpopulation (see Equation 10). This is equivalent to assuming that the time taken for a selective sweep ($\sim \ln[N]/s_b$ generations) is at least as long as the time taken to approach mutation-selection balance ($\sim 1/s_d$ generations). The minimum population size N required for this assumption to hold is of magnitude $\exp[s_b/s_d]$ and as s_d becomes small compared to s_b this rapidly becomes unrealistic. As s_d becomes smaller still, the frequencies of deleterious alleles begin to fluctuate under the influence of genetic drift and our model breaks down in yet another way. Finally when s_d is small compared to the per site mutation rate deleterious alleles reach high frequencies and back mutation can no longer be ignored. There is therefore a need for a more thorough analysis of how fixation probabilities are influenced by very weak selection at very many linked loci in large but not infinite populations.

We have argued that a rapid selective sweep or bottleneck of one individual will purge segregating deleterious alleles from an asexual population and that this will inflate the fixation probabilities of subsequently arising beneficial mutations. In other words, adaptation can trigger further adaptation. This effect would cause selectively driven substitutions to tend to occur in bursts, and the substitution process would tend to be overdispersed relative to a Poisson process. However, the effect seems to be small in many regions of the parameter space. Even when the effect is substantial (*e.g.*, when $s_b = 3 \times 10^{-2}$, $s_d = 10^{-2}$, and $U = 10^{-1}$) our results should not be taken to imply that the net effect would be to speed up adaptation in the long term. This is because the population size of individuals carrying the sweeping beneficial mutation, or the population size during recovery after the bottleneck, will be reduced and less beneficial mutations will arise. Our result shows only that the temporary purging of segregating deleterious alleles will cause a substitution process that is not Poisson in time. [The fixation probabilities of these beneficial alleles will also be increased because they arise in a growing subpopulation (Ewens 1967).] Interestingly, Gerrish (2001) has recently shown that interference between multiple simultaneously segregating beneficial alleles (the Hill-Robertson effect; Fisher 1930; Muller 1932; Hill and Robertson 1966; Felsenstein 1974; Barton 1995; Gerrish and Lenski 1998; Gerrish 2001) will cause the selected substitutions to be "rhythmic," that is, underdispersed or more uniformly spaced in time than random. Determining the interaction between these effects, and whether the net effect is an under- or overdispersed substitution process, requires further work. We note that although Gerrish (2001) performed simulations that explicitly included deleterious alleles, he assumed that $f_0 = 0.9$ at all times, so that only 10% of beneficial alleles arose on a background containing a deleterious allele and nonequilibrium effects were not allowed for at all. We speculate that the index of dispersion for selectively driven substitutions will depend on population size N, since increasing N increases the intensity of the Hill-Robertson effect but has little effect on interference from segregating deleterious alleles. Furthermore, both our argument and the model of Gerrish (2001) assume an unlimited supply of beneficial alleles. More realistic assumptions, such as the mutational landscape model studied by Gillespie (1984), will also cause the substitution process to deviate from a Poisson process. An area for further work is to examine both the effects described here and by Gerrish

(2001) in the context of molecular evolution, where there are a considerable amount of data suggesting that the substitution process is overdispersed (see Cutler 2000 for a review).

It is important to note that our results for nonequilibrium populations consider one of the most extreme nonequilibrium situations possible. Our assumption of a large and completely homogenous population at time $t = 0$ is motivated mostly by mathematical tractability, but it should be a good approximation for some biologically realistic situations, including a population recovering by unchecked binary division after a bottleneck of one individual or a population immediately after a strongly selected ($s_b > 1$) beneficial allele has swept to fixation. In these cases the population size could reach 10^9 in <30 generations, which is rapid compared to the timescale of approach to mutation-selection equilibrium when $s_d \ll 0.03$ (Johnson 1999). Burch and Chao (2000) bottlenecked two populations of the RNA virus $\phi 6$ to size 1 and observed a subsequent approach to mutation-selection equilibrium over a period of ~100 generations. This provides some empirical evidence that nonequilibrium situations such as the one modeled here are relevant. More generally, we hope that our results will motivate further research on evolutionary processes in nonequilibrium populations, under more widely applicable assumptions.

We have shown that the optimal mutation rate, U_{opt}, which is the mutation rate that maximizes the long-term rate of fitness improvement, can depend on the strengths of selection on both deleterious and beneficial alleles. Orr (2000b) previously found that, when selection for beneficial alleles is weaker than against deleterious alleles ($s_b < \min\{s_d\}$), the optimum mutation rate depends only on the strength of selection on deleterious alleles and specifically that $U_{opt} = E_h(s_d)$. Orr (2000b) tentatively suggested that this result might explain Drake's rule, that a range of DNA-based microorganisms with great variation in life history and genome size all have roughly the same per genome per replication mutation rate $\mu_g \simeq 0.003$ (Drake 1991; Drake *et al.* 1998; Sniegowski *et al.* 2000), arguing that it is plausible that $E_h(s_d)$ might be roughly conserved across the group of organisms for which Drake's rule applies. Our results show that this explanation is sensitive to Orr's (2000b) assumption that $s_b < \min\{s_d\}$. Our results suggest that if on the other hand $s_b > s_d$ is typical for microorganisms, then a possible alternative explanation for Drake's rule is that the strength of selection on beneficial alleles could be roughly conserved across the relevant populations. In this respect we note that for the bacterium *E. coli* growing on minimal media Kibota and Lynch (1996) estimated that $\overline{s_d} \leq \sim 0.012$ and that Imhof and Schlötterer (2001) have estimated selection coefficients for beneficial alleles that were often >0.01, conditional on them both becoming fixed and being observed. However, as pointed out by Orr (2000b) and Sniegowski *et al.* (2000), there is only a weak theoretical basis for supposing that asexual populations will evolve a mutation rate that is optimal. Areas for further theoretical research include finding the optimal mutation rate when beneficial and deleterious alleles each have selection coefficients drawn from some distribution of effects and determining under what conditions an asexual population will evolve its optimal mutation rate.

We thank Brian Charlesworth, Arcadi Navarro, Allen Orr, Sally Otto, Mario Pineda-Krch, Rosie Redfield, Olivier Tenaillon, and two anonymous reviewers for discussions and/or helpful comments on the manuscript. T.J. is supported by Wellcome Trust International Prize Travelling Research Fellowship no. 061530. N.B. is supported by the Biotechnology and Biological Sciences Research Council and by the Natural Environment Research Council.

LITERATURE CITED

Barton, N. H., 1987 The probability of establishment of an advantageous mutant in a subdivided population. Genet. Res. **50:** 34–40.

Barton, N. H., 1993 The probability of fixation of a favoured allele in a subdivided population. Genet. Res. **62:** 149–157.

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

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

Barton, N. H., and S. Rouhani, 1987 The frequency of shifts between alternative equilibria. J. Theor. Biol. **125:** 397–418.

Barton, N. H., and M. C. Whitlock, 1997 The evolution of metapopulations, pp. 183–210 in *Metapopulation Biology*, edited by I. A. Hanski and M. E. Gilpin. Academic Press, San Diego.

Berg, O. G., 1995 Periodic selection and hitchhiking in a bacterial population. J. Theor. Biol. **173:** 307–320.

Burch, C. L., and L. Chao, 2000 Evolvability of an RNA virus is determined by its mutational neighbourhood. Nature **406:** 625–628.

Caballero, A., and E. Santiago, 1995 Response to selection from new mutation and effective size of partially inbred populations. 1. Theoretical results. Genet. Res. **66:** 213–225.

Charlesworth, B., 1994 The effect of background selection against deleterious mutations on weakly selected, linked variants. Genet. Res. **63:** 213–227.

Charlesworth, B., 1996 The evolution of chromosomal sex determination and dosage compensation. Curr. Biol. **6:** 149–162.

Cutler, D. J., 2000 Understanding the overdispersed molecular clock. Genetics **154:** 1403–1417.

Davies, E. K., A. D. Peters and P. D. Keightley, 1999 High frequency of cryptic deleterious mutations in *Caenorhabditis elegans*. Science **285:** 1748–1751.

Dawson, K. J., 1999 The dynamics of infinitesimally rare alleles, applied to the evolution of mutation rates and the expression of deleterious mutations. Theor. Popul. Biol. **55:** 1–22.

Drake, J. W., 1991 A constant rate of spontaneous mutation in DNA-based microbes. Proc. Natl. Acad. Sci. USA **88:** 7160–7164.

Drake, J. W., B. Charlesworth, D. Charlesworth and J. F. Crow, 1998 Rates of spontaneous mutation. Genetics **148:** 1667–1686.

Ewens, W. J., 1967 The probability of survival of a new mutant in a fluctuating environment. Heredity **22:** 438–443.

Feller, W., 1968 *An Introduction to Probability Theory and Its Applications*, Vol. I, Ed. 3. John Wiley & Sons, New York.

Feller, W., 1971 *An Introduction to Probability Theory and Its Applications*, Vol. II, Ed. 2. John Wiley & Sons, New York.

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

Fisher, R. A., 1922 On the dominance ratio. Proc. R. Soc. Edinb. **52:** 399–433.

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

Gerrish, P. J., 2001 The rhythm of microbial adaptation. Nature **413:** 299–302.

Gerrish, P. J., and R. E. Lenski, 1998 The fate of competing beneficial mutations in an asexual population. Genetica **102/103:** 127–144.
Gessler, D. D. G., 1995 The constraints of finite size on asexual populations and the rate of the ratchet. Genet. Res. **66:** 241–253.
Gillespie, J. H., 1983a A simple stochastic gene substitution model. Theor. Popul. Biol. **23:** 202–215.
Gillespie, J. H., 1983b Some properties of finite populations experiencing strong selection and weak mutation. Am. Nat. **121:** 691–708.
Gillespie, J. H., 1984 Molecular evolution over the mutational landscape. Evolution **38:** 1116–1129.
Gordo, I., and B. Charlesworth, 2000 The degeneration of asexual haploid populations and the speed of Muller's ratchet. Genetics **154:** 1379–1387.
Haigh, J., 1978 The accumulation of deleterious genes in a population—Muller's ratchet. Theor. Popul. Biol. **14:** 251–267.
Haldane, J. B. S., 1927 A mathematical theory of natural and artificial selection. V. Selection and mutation. Proc. Camb. Philos. Soc. **23:** 838–844.
Harris, T. E., 1963 *The Theory of Branching Processes.* Springer Verlag, Berlin.
Hartl, D. L., and C. H. Taubes, 1996 Compensatory nearly neutral mutations: selection without adaptation. J. Theor. Biol. **182:** 303–309.
Hartl, D. L., and C. H. Taubes, 1998 Towards a theory of evolutionary adaptation. Genetica **102/103:** 525–533.
Hill, W. G., and A. Robertson, 1966 The effect of linkage on limits to artificial selection. Genet. Res. **8:** 269–294.
Imhof, M., and C. Schlötterer, 2001 Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. Proc. Natl. Acad. Sci. USA **98:** 1113–1117.
Johnson, T., 1999 The approach to mutation-selection balance in an infinite asexual population, and the evolution of mutation rates. Proc. R. Soc. Lond. Ser. B **266:** 2389–2397.
Keightley, P. D., 1996 Nature of deleterious mutation load in Drosophila. Genetics **144:** 1993–1999.
Keightley, P. D., and A. Eyre-Walker, 1999 Terumi Mukai and the riddle of deleterious mutation rates. Genetics **153:** 515–523.
Kibota, T. T., and M. Lynch, 1996 Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli.* Nature **381:** 694–696.
Kimura, M., 1957 Some problems of stochastic processes in genetics. Ann. Math. Stat. **28:** 882–901.
Kimura, M., 1983 *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge, UK.
Kimura, M., and T. Maruyama, 1966 The mutational load with epistatic gene interactions in fitness. Genetics **54:** 1337–1351.
Kimura, M., and T. Ohta, 1969 The average number of generations until fixation of a mutant gene in a finite population. Genetics **61:** 763–771.
Kimura, M., and T. Ohta, 1970 Probability of fixation of a mutant gene in a finite population when selective advantage decreases with time. Genetics **65:** 525–535.
Lynch, M., J. Blanchard, D. Houle, T. Kibota, S. Schultz *et al.*, 1999 Perspective: spontaneous deleterious mutation. Evolution **53:** 645–663.
Manning, J. T., and D. J. Thompson, 1984 Muller's ratchet and the accumulation of favourable mutations. Acta Biotheor. **33:** 219–225.
Maruyama, T., 1970 On the fixation probability of mutant genes in a subdivided population. Genet. Res. **15:** 257–260.
Maruyama, T., 1971 An invariant property of a structured population. Genet. Res. **18:** 81–84.
Miralles, R., P. J. Gerrish, A. Moya and S. F. Elena, 1999 Clonal interference and the evolution of RNA viruses. Science **285:** 1745–1747.
Mukai, T., 1964 The genetic structure of natural populations of *Drosophila melanogaster.* I. Spontaneous mutation rate of polygenes controlling viability. Genetics **50:** 1–19.
Muller, H. J., 1932 Some genetic aspects of sex. Am. Nat. **66:** 118–138.
Muller, H. J., 1964 The relation of recombination to mutational advance. Mutat. Res. **1:** 2–9.
Nagylaki, T., 1982 Geographical invariance in population genetics. J. Theor. Biol. **99:** 159–172.
Ohta, T., 1973 Slightly deleterious mutant substitutions in evolution. Nature **246:** 96–98.
Ohta, T., 1992 The nearly neutral theory of molecular evolution. Ann. Rev. Ecol. Syst. **23:** 263–286.
Ohta, T., and J. H. Gillespie, 1996 Development of neutral and nearly neutral theories. Theor. Popul. Biol. **49:** 128–142.
Orr, H. A., 1998 The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution **52:** 935–949.
Orr, H. A., 1999 The evolutionary genetics of adaptation: a simulation study. Genet. Res. **74:** 207–214.
Orr, H. A., 2000a Adaptation and the cost of complexity. Evolution **54:** 13–20.
Orr, H. A., 2000b The rate of adaptation in asexuals. Genetics **155:** 961–968.
Orr, H. A., and Y. Kim, 1998 An adaptive hypothesis for the evolution of the Y chromosome. Genetics **150:** 1693–1698.
Otto, S. P., and M. C. Whitlock, 1997 The probability of fixation in populations of changing size. Genetics **146:** 723–733.
Peck, J. R., 1994 A ruby in the rubbish: beneficial mutations and the evolution of sex. Genetics **137:** 597–606.
Pinkus, A., and S. Zafrany, 1997 *Fourier Series and Integral Transforms.* Cambridge University Press, Cambridge, UK.
Pollak, E., 1966a On the survival of a gene in a subdivided population. J. Appl. Prob. **3:** 142–155.
Pollak, E., 1966b Some effects of fluctuating offspring distributions on the survival of a gene. Biometrika **53:** 391–396.
Pollak, E., 1972 Some effects of two types of migration on the survival of a gene. Biometrics **28:** 385–400.
Poon, A., and S. P. Otto, 2000 Compensating for our load of mutations: freezing the meltdown of small populations. Evolution **54:** 1467–1479.
Rice, W. R., and A. K. Chippindale, 2001 Sexual recombination and the power of natural selection. Science **294:** 555–559.
Robertson, A., 1970 A theory of limits in artificial selection with many linked loci, pp. 246–288 in *Mathematical Topics in Population Genetics,* edited by K.-i. Kojima. Springer-Verlag, Berlin.
Sniegowski, P. D., P. J. Gerrish, T. Johnson and A. Shaver, 2000 The evolution of mutation rates: separating causes and consequences. Bioessays **22:** 1057–1066.
Stephan, W., B. Charlesworth and G. McVean, 1999 The effect of background selection at a single locus on weakly selected, partially linked variants. Genet. Res. **73:** 133–146.
Wolfram, S., 1996 *The Mathematica Book,* Ed. 3. Wolfram Media/Cambridge University Press, Cambridge, UK.

Communicating editor: W. Stephan

APPENDIX A: NONEQUILIBRIUM WILD-TYPE POPULATION

There are many ways in which the wild-type population could be out of equilibrium. For simplicity, we study only one possibility. We assume that the population of interest is descended from one that was free of segregating deleterious alleles at time $t = 0$, where t is measured in generations, and that any deleterious alleles that were fixed at $t = 0$ do not revert to wild-type alleles by back mutation. This initial condition approximates the effect of a rapid selective sweep, a severe population bottleneck with rapid recovery, or the founding of a laboratory evolution experiment from a single clone. As time increases, the frequencies of deleterious alleles will increase and the population will approach an equilibrium at mutation-selection balance. We assume that a single beneficial allele arises at some time $t = \tau > 0$. The analysis in the main part of the article is therefore a special case with $\tau = \infty$. Since the fate of the beneficial

allele depends on the time of its origin, we write $P_{\text{fix},\tau}$ for its fixation probability.

Because we assume that there are no segregating deleterious alleles in the wild-type population at time zero, the mean fitness and distribution of number of deleterious alleles per wild-type individual will change over time. Johnson (1999) gives equations describing these dynamics, in terms of the Laplace transforms (see, *e.g.*, Feller 1971; Pinkus and Zafrany 1997) of the distributions involved. When deleterious alleles have a fixed effect $(1 - s_d)$, the distribution of effects on log fitness has Laplace transform $\tilde{d}(z) = (1 - s_d)^z$, where z is a dummy variable. Equation 12 of Johnson (1999) shows that the Laplace transform of negative log fitness at time t is

$$\tilde{g}(z, t) = \exp\left[U \sum_{i=0}^{t-1} (1 - s_d)^{z+i} - (1 - s_d)^i\right] \quad \text{(A1)}$$

(where g here corresponds to f in the notation of Johnson 1999). This simplifies to

$$\tilde{g}(z, t) = \exp\left[\frac{U(1 - (1 - s_d)^t)}{s_d}((1 - s_d)^z - 1)\right]. \quad \text{(A2)}$$

The population mean fitness at time t, relative to a wild-type mutation-free $(i = 0)$ individual, is given by

$$\bar{w}(t) = \tilde{g}(z = 1, t) = \exp\left[U((1 - s_d)^t - 1)\right]. \quad \text{(A3)}$$

Equation A2 also shows that the distribution of negative log fitness is a scaled Poisson distribution and therefore that the number of deleterious alleles, i, carried by a randomly chosen wild-type individual at time t is Poisson distributed with mean

$$\lambda_t = \frac{U(1 - (1 - s_d)^t)}{s_d}, \quad \text{(A4)}$$

where the quantities in Equations A3 and A4 are given after mutation and before selection. Equation A4 allows us to calculate the probability of the beneficial mutation arising in a given background at time $t = \tau$, and Equation A3 allows us to calculate the absolute fitnesses of individuals carrying the beneficial allele for all subsequent times $t \geq \tau$. As expected, as $t \to \infty$ these quantities approach their equilibrium values $\bar{w} \to e^{-U}$ (Kimura and Maruyama 1966) and $\lambda \to U/s_d$ (Haigh 1978).

Equations A3 and A4 can also be derived using factorial cumulant-generating functions, as described by Dawson (1999).

The fixation probability of a beneficial mutation arising at time $t = \tau$ can be obtained by choosing some large time $t = T$, setting $P_{i,T} = P_{i,\infty}$, and then finding the $P_{i,t}$ recursively in decreasing order of t using Equation 9 with the upper limit in the sum set to $i_{\max}$. Results obtained in this way, using $T = 1000$, are shown in Figures 3 and 4. The choice of $T = 1000$ is justified by noting that $\bar{w}_t$ and λ_t approach their equilibria over a timescale proportional to $1/s_d = 100$ generations. Similar calculations for smaller s_d would be time-consuming at present because both the appropriate value for t and the number of genetic backgrounds to be considered for a given s_b ($i_{\max}$) scale with $1/s_d$ and hence the number of calculations scales with $1/s_d^2$.

Given the computationally intensive nature of these calculations, an approximation is desirable. It is tempting to assume that $\bar{w}_t$ changes slowly relative to the time taken for the fate of the branching process to be determined (which is typically "rapid"; see, *e.g.*, Feller 1968, p. 297), which would give $P_{i,t+1} \simeq P_{i,t} \simeq P_{i,\tau}$. Then the method used for $\tau = \infty$ could be followed with appropriately modified values for the W_i and f_i. Numerical results (not shown) show that this approximation is highly unsatisfactory for many combinations of parameter values. Specifically, this approximation often gives $P_{\text{fix},\tau} < P_{\text{fix},\infty}$ when more accurate calculations show that $P_{\text{fix},\tau} > P_{\text{fix},\infty}$. Because this approximation assumes that the fate of the beneficial allele is determined rapidly relative to the mean fitness dynamics in the wild-type subpopulation, it is most accurate for large s_b and/or large τ. However, sufficiently large s_b are so large that there is negligible effect from interfering deleterious alleles and sufficiently large τ are so large that solutions are indistinguishable from solutions for equilibrium wild-type populations.

APPENDIX B: NONEQUILIBRIUM POPULATION AND BENEFICIAL ALLELE OF SMALL EFFECT

For the special case $s_b \leq s_d$ we can prove an important result. This is that the net fixation probability of a beneficial allele arising in a nonequilibrium population (for the specific departure from equilibrium considered in appendix a) is greater than in an equilibrium population. More specifically, going backward in time from a close-to-equilibrium population, the net fixation probability increases. We speculate that this result is probably true in general but do not see an easy way to prove it. We can also show that for $s_b \leq s_d$ the net fixation probability increases by at most a factor $(1 + s_b)$ per generation backward in time. This may explain why beneficial alleles of small effect do not have inflated fixation probabilities in nonequilibrium populations.

When $s_b \leq s_d$ it is necessary only to consider the fate of beneficial alleles in a background of zero deleterious alleles; that is, $i_{\max} = 0$. The recursion for the fixation probability of a beneficial allele on such a background is

$$P_{0,t} = 1 - \exp\left[-(1 + s_b)\frac{e^{-U}}{\bar{w}_t}P_{0,t+1}\right]. \quad \text{(B1)}$$

The net fixation probability of a beneficial allele that arises at time $t = \tau$ is

$$P_{\text{fix},\tau} = f_{0,\tau}P_{0,\tau}, \quad \text{(B2)}$$

where $f_{0,t} = \exp[-\lambda_t]$ is the probability of arising in a background free of deleterious alleles, at time t. Equation A4 shows that

$$\frac{f_{0,t+1}}{f_{0,t}} = \frac{e^{-U}}{\overline{w}_t}. \tag{B3}$$

By substituting Equations B2 and B3 into Equation B1 we can show that

$$P_{\text{fix},\tau} = f_{0,\tau}\left(1 - \exp\left[-\frac{(1 + s_b)}{f_{0,\tau}} P_{\text{fix},\tau+1}\right]\right). \tag{B4}$$

The series expansion

$$P_{\text{fix},\tau} = (1 + s_b)P_{\text{fix},\tau+1} - \frac{(1 + s_b)^2}{2f_{0,\tau}} P^2_{\text{fix},\tau+1} + \frac{(1 + s_b)^3}{6f^2_{0,\tau}} P^3_{\text{fix},\tau+1} - \ldots \tag{B5}$$

converges for all $f_{0,\tau} > 0$ and shows that

$$P_{\text{fix},\tau} \leq (1 + s_b)P_{\text{fix},\tau+1} \tag{B6}$$

as claimed.

Our proof that net fixation probabilities increase as we go backward in time seems a little contorted, but we have been unable to find a more transparent proof. We first assume that a beneficial allele that arises in a nonequilibrium population at some large time $\tau = T$ has net fixation probability equal to that for an equilibrium population

$$P_{\text{fix},T} = P_{\text{fix},\infty} \tag{B7}$$

and then prove by induction that

$$P_{\text{fix},0} < P_{\text{fix},\tau} < P_{\text{fix},\tau+1} < P_{\text{fix},T} \tag{B8}$$

for all $0 < \tau < T - 2$. As $T \to \infty$ condition (B7) is necessarily satisfied and therefore

$$P_{\text{fix},0} < P_{\text{fix},\tau} < P_{\text{fix},\tau+1} < P_{\text{fix},\infty} \tag{B9}$$

for all $0 < \tau$.

To prove Equation B8, we note first that, from (B4), the equation

$$P_{\text{fix},\tau} = P_{\text{fix},\tau+1} \tag{B10}$$

is satisfied only when $P_{\text{fix},\tau} = 0$ or $P_{\text{fix},\tau} = f_{0,t}p[s_b]$. Therefore a graph of $P_{\text{fix},\tau}$ against $P_{\text{fix},\tau+1}$ does not touch the line $P_{\text{fix},\tau} = P_{\text{fix},\tau+1}$ between these two points, and since it has gradient $(1 + s_b)$ at the origin (see Equation B5) it must in fact lie above the line. Therefore

$$P_{\text{fix},\tau} > P_{\text{fix},\tau+1} \quad \text{when } 0 < P_{\text{fix},\tau+1} < f_{0,\tau}p[s_b]. \tag{B11}$$

Noting that $P_{\text{fix},t}$ is a monotonically increasing function of $P_{\text{fix},t+1}$ (see Equation B4) and since $f_{0,\tau-1} > f_{0,\tau}$ we have

$$P_{\text{fix},\tau} < f_{0,\tau}p[s_b] < f_{0,\tau-1}p[s_b] \quad \text{when } 0 < P_{\text{fix},\tau+1} < f_{0,\tau}p[s_b]. \tag{B12}$$

Observing that the condition in Equations B11 and B12 is satisfied when $\tau + 1 = T$ because by assumption $P_{\text{fix},T} = P_{\text{fix},\infty} = f_{0,\infty}p[s_b]$, Equations B11 and B12 constitute a proof by induction for Equation B8.

APPENDIX C: APPROXIMATIONS

In the analysis section of the article we described a method that allows us to calculate P_{fix} exactly (given the model assumptions) for any parameter values. However, this analysis makes no real improvement to our understanding of how fixation probabilities are influenced by segregating deleterious alleles. There does not appear to be a more concise exact expression for P_{fix}, except for the special case where $s_b < s_d$ and hence $i_{\max} = 0$, which was solved by Peck (1994).

In this section we assume an equilibrium wild-type population, but the methods used could in principle be applied to nonequilibrium wild-type populations.

It is useful to write $P_i = p_i + x_i$. Here the first term gives the probability that the beneficial allele fixes within the genetic background in which it arose, and the second term gives the probability, conditional on it being lost from that background and that it was lost by mutation, that it fixes in some lower fitness background where it still confers a net advantage. $x_{i_{\max}} = 0$ but Manning and Thompson's (1984) analysis erroneously assumed $x_i = 0$ for all i. By making this decomposition of P_i, separating out the $j = 0$ term from the sum in Equation 12 and then rearranging we can obtain

$$p_i + x_i = (1 - \exp[-e^{-U}W_i p_i]) + \exp[-e^{-U}W_i p_i] \times \left(1 - \exp\left[-e^{-U}W_i\left(x_i + \sum_{j=1}^{i_{\max}-i} \frac{U^j}{j!} P_{i+j}\right)\right]\right), \tag{C1}$$

which by using the identity $p_i = 1 - \exp[-e^{-U}W_i p_i]$ can be simplified to

$$x_i = (1 - p_i) \times \left(1 - \exp\left[-e^{-U}W_i\left(x_i + \sum_{j=1}^{i_{\max}-i} \frac{U^j}{j!} P_{i+j}\right)\right]\right). \tag{C2}$$

When selection and mutation are weak, we can assume $1 - p_i \simeq 1$, $e^{-U}W_i \simeq 1$, and ignore terms in U^2 and higher powers in the summation to obtain

$$x_i \simeq 1 - \exp[-x_i - UP_{i+1}], \tag{C3}$$

which can be solved by making a series expansion to obtain the quadratic equation

$$x_i \simeq x_i + UP_{i+1} - \frac{1}{2}\left(x_i^2 + 2x_iUP_{i+1} + U^2P_{i+1}^2\right)$$

$$0 \simeq x_i^2 + x_i 2UP_{i+1} + U^2P_{i+1}^2 + 2UP_{i+1} \tag{C4}$$

and hence

$$x_i \simeq -UP_{i+1} + \sqrt{2UP_{i+1} - (UP_{i+1})^2}, \tag{C5}$$

which for $UP_{i+1} \ll 1$ gives $x_i \simeq \sqrt{2UP_{i+1}}$ and so

$$P_i \simeq p_i + \sqrt{2UP_{i+1}}. \tag{C6}$$

Unfortunately, even this simple expression does not seem to lead to an approximation for P_i in terms of the

model parameters alone (*i.e.*, which does not depend on P_{i+1}). Such an approximation can be found only when selection is weak relative to mutation, so that $p_i \ll \sqrt{2UP_{i+1}}$, and then we have

$$P_i \simeq P_{i_{\max}}^{\ A}\,(2U)^{1-A}, \tag{C7}$$

where $A = 1/(2^{i_{\max}-i})$. This shows that $P_{i_{\max}-i} \to 2U$ as i increases (see Figure 1). This approximation may be useful more generally, so long as $p_i \ll \sqrt{2UP_{i+1}}$ for the backgrounds that contribute significantly to P_{fix}. However, we do not see a way to sum this approximation for the P_i over all backgrounds and hence do not see a way toward an approximate formula for P_{fix}.

We can use Equation C6 to show that when $s_b \gg s_d$ the relative fixation probability $R = P_{\text{fix}}/p[s_b]$ depends only on the compound parameters $\sigma_b = s_b/U$ and $\sigma_d = s_d/U$. First note that under these conditions $i_{\max} \simeq s_b/s_d = \sigma_b/\sigma_d$ and that the f_i depend only on $\lambda = U/s_d = 1/\sigma_d$. Let the contribution to R from background i be $R_i = P_i/p[s_b] \simeq P_i/2s_b$. Then from Equation C6, which assumes weak selection and weak mutation, we have

$$R_i \simeq \frac{2(s_b - is_d)}{2s_b} + \sqrt{\frac{2U}{2s_b}\frac{P_{i+1}}{2s_b}}, \tag{C8}$$

which rearranges to give

$$R_i \simeq 1 - i\frac{\sigma_d}{\sigma_b} + \sqrt{\frac{1}{\sigma_b}R_{i+1}}. \tag{C9}$$

Although $R_{i_{\max}}$ cannot be written in terms of the compound parameters σ_b and σ_d alone, the R_i rapidly become independent of $R_{i_{\max}}$ as i decreases. Therefore, to a good approximation R depends only on σ_b and σ_d.