Genet. Res., Camb. (1999), **73**, *pp.* 133–146. *With 6 figures.* Printed in the United Kingdom © 1999 Cambridge University Press

The effect of background selection at a single locus on weakly selected, partially linked variants

WOLFGANG STEPHAN[1]*, BRIAN CHARLESWORTH[2] AND GILEAN McVEAN[2]
[1] *Department of Biology, University of Rochester, Rochester, NY 14627-0211, USA*
[2] *Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, UK*

(*Received 18 May 1998 and in revised form 18 September 1998*)

Summary

Previous work has shown that genetic diversity at a neutral locus is affected by background selection due to recurrent deleterious mutations as though the effective population size N_e is reduced by a factor that is calculable from genetic parameters such as mutation rates, selection coefficients, and the rates of recombination between sites subject to selection and the neutral locus. Given that silent changes at third coding positions are often subject to weak selection pressures, it is important to develop similar quantitative predictions of the effects of background selection on variation and evolution at weakly selected sites. A diffusion approximation is derived that describes the effects of the presence of a single locus subject to mutation and strongly deleterious selection on variation and evolution at a partially linked, weakly selected locus. The results are validated by computer simulations using the Ito pseudo-sampling method. We show that both nucleotide site diversity and rates of molecular evolution at a weakly selected locus are affected by background selection as though N_e is reduced in the same way as for a neutral locus. Heuristic arguments are presented as to why the change in N_e for the neutral case also applies with weak selection. As in the case of a neutral locus, the number of segregating sites in the population is poorly predicted from the change in N_e. The potential significance of the results in relation to the effects of recombinational environment on molecular variation and evolution is discussed.

1. Introduction

There has recently been a great deal of theoretical work on variation and evolution at loci that are closely linked to sites that are the targets of selection. This work has largely been motivated by the observation of reduced DNA variability in natural populations of *Drosophila* at loci situated in regions where genetic recombination is relatively infrequent, compared with regions where it occurs at normal frequencies (reviewed by Aguadé & Langley, 1994; Aquadro *et al.*, 1994; Moriyama & Powell, 1996). Similar patterns have recently been detected in *Mus* (Nachman, 1997), *Aegilops* (Dvorak *et al.*, 1998) and *Lycopersicon* (Stephan & Langley, 1998). In addition, the level of codon bias in *D. melanogaster* is lower in regions of reduced recombination, suggesting that selection at weakly selected sites is less effective when recombination is infrequent (Kliman & Hey, 1993). These observations can be explained by (i) 'selective sweeps' of favourable mutations, which result in the fixation of adjacent chromosomal regions (Maynard Smith & Haigh, 1974; Kaplan *et al.*, 1989; Stephan *et al.*, 1992; Stephan, 1995), (ii) 'background selection', in which neutral or nearly neutral variants are lost as a result of linkage to strongly deleterious mutant alleles that are destined to be rapidly eliminated from the population (Charlesworth *et al.*, 1993, 1995; Charlesworth, 1994, 1996; Hudson, 1994; Hudson & Kaplan, 1994, 1995; Nordborg *et al.*, 1996), (iii) temporally varying selection pressures, which can cause linked variants to be brought close to loss or fixation (Gillespie, 1994, 1997; Barton, 1995).

In this paper we will be concerned solely with the effects of background selection. Previous work on this problem has concentrated on the properties of neutral

* Corresponding author. Telephone: +1 (716) 275 009. Fax: +1 (716) 275 2070. e-mail: stephan@troi.cc.rochester.edu.

variants linked to loci under selection. The reduction in genetic diversity at neutral sites predicted by the background selection model can be largely understood in terms of a reduction in effective population size (N_e) caused by the fitness effects of the loci under selection (Charlesworth *et al.*, 1993), and useful formulae for this reduction have been developed (Hudson, 1994; Hudson & Kaplan, 1994, 1995; Nordborg *et al.*, 1996; Santiago & Caballero, 1998). Charlesworth (1994) derived results for the effects of background selection on the probabilities of fixation of weakly selected favourable or deleterious alleles in a non-recombining genome; Peck (1994) studied the effect of background selection on the fixation probability of a strongly selected favourable allele in a non-recombining genome subject to recurrent deleterious mutations. Barton (1995) has obtained results for the effect of recurrent deleterious mutations at a single locus on the fixation probability of a linked, strongly selected, favourable mutation.

The results of these calculations suggest that background selection affects fixation probabilities through a reduction in N_e, in a way which is virtually identical to its effect on diversity in the neutral case. But no theoretical results on fixation probabilities or genetic diversity for the effect of background selection on partially linked, weakly selected loci are available. (By weak selection, we mean that the product of the effective population size and the selection coefficient is of the order of one, so that finite population size effects can significantly influence the fate of the alleles in question.) Given the evidence that such weak selection pressures are important in controlling patterns of variation and evolution at synonymous sites in *Drosophila* (Akashi, 1995, 1996; Akashi & Schaeffer, 1997), it is clearly important to have a quantitative theory of the effects of background selection on weakly selected sites. The purpose of the present paper is to present some analytical and simulation results for the effects of the presence of a single locus subject to mutation and strongly deleterious selection on a partially linked, weakly selected locus. We show that both nucleotide site diversity and rates of molecular evolution at the weakly selected locus are affected by background selection as though N_e is reduced in the same way as for a neutral locus.

2. The model and its assumptions

We consider a two-locus, two-allele system, in a population of N diploid, randomly mating individuals. We assume that N is sufficiently large and selection is sufficiently strong that the alleles A and a at the first locus are in an approximate equilibrium, due to selection against the deleterious mutant allele a. Following the notation of Nordborg *et al.* (1996), the frequency of the wild-type allele A is denoted by p and that of the mutant allele a by q. Furthermore, we assume that $q \ll 1$; thus, selection acts primarily against heterozygous carriers of the mutant allele. Let u be the mutation rate from A to a, and t be the product of the dominance coefficient and the coefficient of selection against mutant homozygotes; then $q = u/t$ (approximately), according to the classical formula for equilibrium under mutation and selection. The second locus is assumed to be under directional selection; furthermore, it is assumed that dominance is intermediate, so that Bb and BB individuals have fitness $1+s$ and $1+2s$, respectively, relative to a fitness of 1 for bb. (The latter assumption could be generalized to arbitrary degrees of dominance without changing the main conclusions of this paper.) s may be negative or positive. We assume that $|s| \ll t$. Recombination between the two loci is measured by the recombination fraction, r'.

Since the first locus is assumed to be in equilibrium, it is convenient to deviate from the standard notation of two-locus population genetics theory and to introduce the following variables: x is the frequency of B among chromosomes carrying allele A, and y is the frequency of B among chromosomes carrying allele a. Then the deterministic recurrence equations for the changes in x and y are (Nordborg *et al.*, 1996)

$$\left.\begin{aligned}\Delta x &= sx(1-x)-qr(x-y),\\ \Delta y &= sy(1-y)+p(r+t)(x-y),\end{aligned}\right\} \tag{1}$$

where $r = r'(1-t)$. These equations are exact to $O(s)$ and $O(q)$. On the boundary, the vector field (defined by the right-hand side of (1)) is directed towards the interior of the unit square, except at $(0,0)$ and $(1,1)$.

In parallel, we consider the corresponding diffusion process. On the time scale of generations (measured by τ), the Kolmogorov forward equation for the probability densities of x and y is (cf. Nordborg *et al.*, 1996)

$$\frac{\partial}{\partial\tau}p(x,y,\tau) = Lp(x,y,\tau), \tag{2}$$

where

$$L = \frac{1}{4Nq}\frac{\partial^2}{\partial y^2}y(1-y)-\frac{\partial}{\partial y}[sy(1-y)+p(r+t)(x-y)]$$
$$+\frac{1}{4Np}\frac{\partial^2}{\partial x^2}x(1-x)-\frac{\partial}{\partial x}[sx(1-x)-qr(x-y)]. \tag{3}$$

Equation (2), with L defined as in (3), is exact to $O(N^{-1}, s)$; quadratic and higher-order terms in q are also neglected in the drift operator. To define the diffusion problem completely, initial and boundary conditions have to be specified. With respect to the initial value, we assume throughout this paper that a diffusion process which is at (x,y) at time τ has started

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

at (ξ, η) at time 0. The definition of boundary conditions is less clear, because a general theory of boundary conditions is not available for two-dimensional diffusions. However, it appears that no conditions can be imposed on the boundary of the unit square, except for $(0,0)$ and $(1,1)$. This is because the boundary of the unit square is not invariant under the vector field defined by the right-hand side of (1), except that $(0,0)$ is mapped onto itself, as is $(1,1)$. The diffusion will eventually be absorbed at $(0,0)$ or $(1,1)$. Therefore, we will be able to impose conditions on the probability of ultimate fixation of allele B, which we calculate in Section 4.

3. Approximate solution of the diffusion equation

The diffusion equation (2) cannot be solved exactly by analytical methods. In this section, we apply an approximation method that allows us to reduce the diffusion equation to an exactly soluble one. To describe this procedure, we start with the deterministic equations (1). Because we assumed that $q \ll 1$ and $|s| \ll t$, the dynamics of the deterministic system are such that the trajectory, starting at some point (ξ, η), rapidly approaches a quasi-equilibrium state $y = y(x)$ near the diagonal $y = x$. After this rapid relaxation phase, the system moves slowly towards the stationary solution (i.e. $(0,0)$ or $(1,1)$), while staying close to the diagonal. The fast-relaxing variable is y, because for initial values $\xi \neq \eta$ its dynamics are mainly determined by the parameter $p(r+t)$; because of our standard assumptions ($q \ll 1$ and $|s| \ll t$), this parameter is much larger than the other two parameters of (1), s and q. This procedure thus leads to an 'adiabatic' elimination of variable y, which varies rapidly on the time scale of the slowly varying variable x. The two equations can therefore be decoupled and solved successively, starting with the equation for y.

Similar elimination procedures have been developed for diffusion processes. We follow the method of Gardiner (1990, chap. 6.6.3). This method leads to more accurate results than the standard approximation procedure used in population genetics for the elimination of fast-changing variables (which are in a quasi-equilibrium) from multidimensional diffusion equations (e.g. Kimura, 1985; Stephan, 1996). We first write the diffusion operator L as

$$L = bL_1 + L_2 + L_3 \tag{4a}$$

where $b = p(r+t)$ is the relaxation coefficient of the fast variable, y, and L_1 describes the diffusion of the fast variable y (which depends on x); thus,

$$L_1 = \frac{1}{4Nqb}\frac{\partial^2}{\partial y^2} y(1-y) - \frac{\partial}{\partial y}\left[\frac{s}{b} y(1-y) + x - y\right]. \tag{4b}$$

Furthermore, the exchange between the 'subpopulations' of A and a chromosomes (caused by recombination and mutation) and the slow movement along the diagonal $y = x$ (caused by selection) will be decoupled as follows:

$$L_2 = -a\frac{\partial}{\partial x}(y - y(x)), \tag{4c}$$

and

$$L_3 = \frac{1}{4Np}\frac{\partial^2}{\partial x^2} x(1-x) - \frac{\partial}{\partial x}[sx(1-x) - a(x - y(x))], \tag{4d}$$

where $a = qr$. We also write $p_x(y)$ for the stationary solution of the fast dynamics; i.e. the solution of

$$L_1 p_x(y) = 0, \tag{5a}$$

and the quasi-equilibrium for y is given by

$$y(x) = \int_0^1 y p_x(y)\, dy. \tag{5b}$$

Gardiner (1990) has shown that an approximate solution of (2) and (3) can be obtained in terms of the Laplace transform. The Laplace transform of any function of time $f(\tau)$ is defined by

$$\tilde{f}(\sigma) = \int_0^\infty \mathrm{e}^{-\sigma\tau} f(\tau)\, d\tau. \tag{6}$$

The solution, which is accurate to $O(N^{-1}, s)$ is then found as (Gardiner, 1990, eqs. [6.6.83] and [6.6.105])

$$\sigma\tilde{\nu}(\sigma) = \left\{PL_3 - \frac{1}{b} PL_2 L_1^{-1}(L_2 + [L_3, P])\right\}\tilde{\nu}(\sigma) + \nu(0). \tag{7}$$

Here we have used the projection operator

$$Pf(x,y) = p_x(y)\int_0^1 f(x,y)\, dy, \tag{8}$$

where $f(x,y)$ is an arbitrary function. $[L_3, P]$ is a commutator defined as $[L_3, P] = L_3 P - PL_3$. It follows from (8) that

$$\nu(\tau) = p_x(y)\,\hat{p}(x,\tau), \tag{9a}$$

where

$$\hat{p}(x,\tau) = \int_0^1 p(x,y,\tau)\, dy \tag{9b}$$

and $p(x,y,\tau)$ is the solution of (2) and (3).

We first consider the term $PL_3\,\tilde{\nu}(\sigma)$ of (7). Using (9), we find

$$PL_3\,\nu(\tau) = p_x(y)\left\{\frac{1}{4Np}\frac{\partial^2}{\partial x^2} x(1-x) - \frac{\partial}{\partial x} \times [sx(1-x) - a(x - y(x))]\right\}\hat{p}(x,\tau). \tag{10}$$

The term in curly brackets is the infinitesimal operator of a diffusion in x which would result if the fast-changing variable y were eliminated – as is generally done in population genetic applications of diffusion

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

theory – by simply using the deterministic equations (1). However, as simulation has revealed (results not shown), this reduction procedure produces poor results in this case. We therefore evaluate the remaining terms of (7).

We begin with the operator $PL_2L_1^{-1}L_2$. A convenient expression can be found if we use the identity (Gardiner, 1990, eq. [6.5.33])

$$L_1^{-1}\Big(1-P\Big) = -\int_0^\infty \mathrm{e}^{L_1\tau}\,d\tau. \tag{11}$$

Using (11), we immediately find

$$PL_2L_1^{-1}L_2p_x(y)\hat{p}(x) = -p_x(y)\int_0^1 dy'\,L_2 \times \int_0^\infty \mathrm{e}^{L_1\tau}\,d\tau\,L_2p_x(y')\hat{p}(x). \tag{12}$$

To evaluate the integrals, we need an explicit expression for $p_x(y)$, which is defined by (5*a*). Equation (5*a*) can be written in the form

$$\frac{1}{2}\frac{\partial^2}{\partial y^2}y(1-y)p_x(y) - \frac{\partial}{\partial y}[\alpha y(1-y) - \beta(1-x)y + \beta x(1-y)]p_x(y) = 0, \tag{13}$$

where we use the abbreviations $\alpha = 2Nqs$ and $\beta = 2Nqb$. The form of (13) is well known (e.g. Ewens, 1979, chap. 5.6); its solution is

$$p_x(y) = C_x y^{2\beta x-1}(1-y)^{2\beta(1-x)-1}\,\mathrm{e}^{2\alpha y}, \tag{14}$$

where C_x is a normalization constant. Furthermore, we need the partial derivative

$$\frac{\partial}{\partial x}p_x(y) = p_x(y)\left[\ln\frac{y}{1-y} - E_{ss}\left(\ln\frac{y}{1-y}\right)\right], \tag{15a}$$

where $E_{ss}(\ldots)$ denotes the expectation with regard to the stationary solution $p_x(y)$. Since the diffusion process is expected to stay close to the quasi-equilibrium $y = y(x)$ after the short initial phase, we may expand the function in the square brackets into a Taylor series about $y = y(x)$:

$$\frac{\partial}{\partial x}p_x(y) = 2\beta p_x(y) \times \left\{\frac{1}{y(x)(1-y(x))}[y-y(x)] - \frac{1}{2}\frac{1-2y(x)}{(y(x)(1-y(x))^2} \times \left[(y-y(x))^2 - E_{ss}((y-y(x)^2))\right] + O(y-y(x))^3\right\} \tag{15b}$$

Thus, using (15*b*) with terms up to second order in $y-y(x)$, we can write

$$PL_2L_1^{-1}L_2p_x(y)\hat{p}(x) = -a^2p_x(y)\frac{\partial}{\partial x}\left\{-D_0(x)\frac{dy(x)}{dx} + D_1(x) \times \left[\frac{\partial}{\partial x} + \beta\frac{1-2y(x)}{(y(x)(1-y(x))^2}E_{ss}((y-y(x)^2)\right] + 2\beta D_2(x) \times \frac{1}{y(x)(1-y(x))} - \beta D_3(x)\frac{1-2y(x)}{(y(x)(1-y(x))^2}\right\}\hat{p}(x), \tag{16}$$

where

$$D_n(x) = \int_0^1 dy'(y'-y(x)) \times \int_0^\infty \mathrm{e}^{L_1\tau}\,d\tau(y'-y(x))^n\,p_x(y'), \quad n \geqslant 0. \tag{17}$$

$D_n(x)$ can be computed in a straightforward way (Appendix, Section (i))

$$D_n(x) = \left[1 + \frac{s}{b}(1-2y(x))\right]E_{ss}((y-y(x))^{n+1}), \tag{18}$$

which is $O(s)$. The stationary moments can also be calculated to a sufficient degree of accuracy (Appendix, Section (ii)). Using the formulae for the moments (A.7–A.10), (16) leads to

$$-\frac{1}{b}PL_2L_1^{-1}L_2\nu(\tau) = p_x(y)\left\{\frac{1}{4N}q\frac{r^2}{(r+t)^2}\frac{\partial^2}{\partial x^2}\right\}\hat{p}(x,\tau). \tag{19}$$

This equation is $O(N^{-1}, s)$. Quadratic terms in q have also been neglected.

Finally, we have to evaluate the operator $PL_2L_1^{-1}[L_3, P]$, where (Gardiner, 1990, eq. [6.6.106]) we have

$$[L_3, P]\nu(\tau) = p_x(y)\left\{r_x(y) \times \left[\frac{1}{4Np}\frac{\partial}{\partial x}x(1-x) - [sx(1-x) - a(x-y(x))\right] + s_x(y)\frac{1}{4Np}x(1-x)\right\}\hat{p}(x,\tau), \tag{20a}$$

and

$$r_x(y) = \frac{(\partial p_x(y))/\partial x}{p_x(y)}, \quad s_x(y) = \frac{(\partial^2 p_x(y))/\partial x^2}{p_x(y)}. \tag{20b}$$

A straightforward calculation yields

$$\frac{1}{b}PL_2L_1^{-1}[L_3, P]\nu(\tau) = p_x(y)\left\{-\frac{1}{2N}q\frac{r}{r+t}\frac{\partial^2}{\partial x^2}x(1-x) + sq\frac{r}{r+t}\frac{\partial}{\partial x}(x(1-x)\right\}\hat{p}(x,\tau). \tag{21}$$

Adding (10), (19) and (21), and using (7), results in a

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

diffusion equation that describes the dynamics of the slow variable x:

$$\frac{\partial}{\partial\tau}\hat{p}(x,\tau)=\frac{1}{4N}\left(1+q\frac{t^2}{(r+t)^2}\right)\frac{\partial^2}{\partial x^2}x(1-x)\times\hat{p}(x,\tau)-s\frac{\partial}{\partial x}x(1-x)\hat{p}(x,\tau). \quad (22)$$

Again, this equation is $O(N^{-1},s)$, and neglects quadratic and higher-order terms in q. It has the same form as a one-dimensional diffusion equation with directional selection. The additional factor (in front of the diffusion operator) indicates that the effective population size is reduced by $1-qt^2/(r+t)^2$. The same reduction in N_e has been found by Hudson & Kaplan (1994, 1995) and Nordborg *et al.* (1996) in their analysis of the effects of background selection and recombination on neutral variation.

4. Applications of the results

(i) *Fixation probability and the rate of molecular evolution*

Equation (22) shows that the distribution of allele frequencies among the A-carrying chromosomes is described by the standard forward diffusion equation of population genetics (Kimura, 1964, 1983), with a simple modification to the effective population size, N_e, which describes the sampling variance of allele frequency over a single generation. Since the forward equation can, in principle, be used to obtain all properties of interest, including fixation probabilities and sojourn times in specified intervals of gene frequency (Kimura, 1964, 1983), this result can be used to obtain the relevant formulae, by substituting this expression for N_e into standard equations.

We first provide an approximate formula for the fixation probability of a weakly selected mutant. To a good approximation, it should be legitimate to consider only the distribution of B among A chromosomes, as in (22), since (1) imply that fixation of B in this class essentially guarantees fixation in the population as a whole, and is certainly required for fixation. Using standard results (Kimura, 1964, 1983), the probability of fixation, $U(\xi)$, of a mutant allele B with initial frequency ξ becomes

$$U(\xi)=\frac{1-e^{-4Nfs\xi}}{1-e^{-4Nfs}}, \quad (23a)$$

where

$$f=1-q\frac{t^2}{(r+t)^2}. \quad (23b)$$

For large Nfs, (23*a*) for the case of a favourable mutation present as a single copy ($\xi=1/2N$) approaches Barton's (1995) (17*c*):

$$U\left(\frac{1}{2N}\right)=2sf. \quad (23c)$$

On the assumption that long-term evolutionary change results from the fixation of new mutations that occur as unique events, the rate of molecular evolution, k, is given by the rate of input of new mutations into the population, multiplied by their probability of fixation (Kimura, 1983, chap. 3). Thus, in the present case we have

$$k=2N\mu\,U\left(\frac{1}{2N}\right), \quad (24)$$

where μ denotes the mutation rate with respect to the segment of the genome under consideration.

(ii) *Sojourn times and nucleotide site diversity*

Using (23) and (4.24) from Ewens (1979), we immediately obtain a formula for the sojourn time density for our one-dimensional diffusion, which describes the expected time that the allele B with initial frequency ξ spends between frequencies x and $x+dx$ among the chromosomes carrying A:

$$t(x,\xi)=U(\xi)\frac{1-e^{4Nfs(x-1)}}{sx(1-x)}, \quad \xi\leqslant x\leqslant 1. \quad (25)$$

As discussed by Charlesworth *et al.* (1993) and Charlesworth (1994), under the infinite sites model with unidirectional mutations from b to B arising in each generation, the average nucleotide diversity, π, within the A class contributed by such variants, relative to the classical neutral value, π_0, is given by

$$\frac{\pi}{\pi_0}=\tfrac{1}{2}H, \quad (26a)$$

where

$$H=\int_0^1 2x(1-x)\,t(x,\xi)\,dx=2\frac{U(\xi)}{s}\left(1-\frac{1-e^{-4Nfs}}{4Nfs}\right). \quad (26b)$$

For $s\to 0$, H converges to $4Nf\xi$; with $\xi=1/2N$, we obtain the familiar result $\pi/\pi_0=f$ for the effect of background selection on neutral variants (Hudson & Kaplan 1994, 1995; Nordborg *et al.*, 1996).

If q is small, as we have assumed, the overall diversity in the population is overwhelmingly determined by the diversity within the A class (Nordborg

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

et al., 1996), and so (26) should provide an excellent approximation for the diversity in the population as a whole; i.e. Nf can be used to replace N_e in the standard equation for diversity at weakly selected sites (Kimura, 1983, p. 45; Charlesworth, 1994). We would not, however, expect this to be true of the number of segregating sites maintained in the population under drift–selection–mutation equilibrium, since this is proportional to the expected sojourn time of a variant in the whole population (Ewens, 1979, p. 238). Unless the population size is extremely large, the time to loss of a B mutation that is associated with a is not greatly different from that for a neutral allele, so that the presence of a alleles does not have as marked an effect on the persistence of B variants as on the net diversity, to which a contributes little (Charlesworth *et al.*, 1993). This applies to both neutral and weakly selected B variants.

5. Simulation results

(i) *Simulation methods*

In order to examine the validity of the analytical approximations derived above, we have performed simulations of the fate of novel mutants in small populations ($N = 2000$ diploid individuals). The above system of two loci with two alleles was simulated. At the A locus, a constant frequency of the strongly deleterious allele a (with heterozygous selection coefficient $t = 0{\cdot}02$) was maintained at the deterministic equilibrium of $q = 0{\cdot}1$ (where $q = u/t$), while novel mutations (b to B) of intermediate dominance ($h = 1/2$) were introduced as single copies at the B locus, separated from the first by recombination fraction r'. The effects of selection and drift on variants on chromosomes carrying A are independent

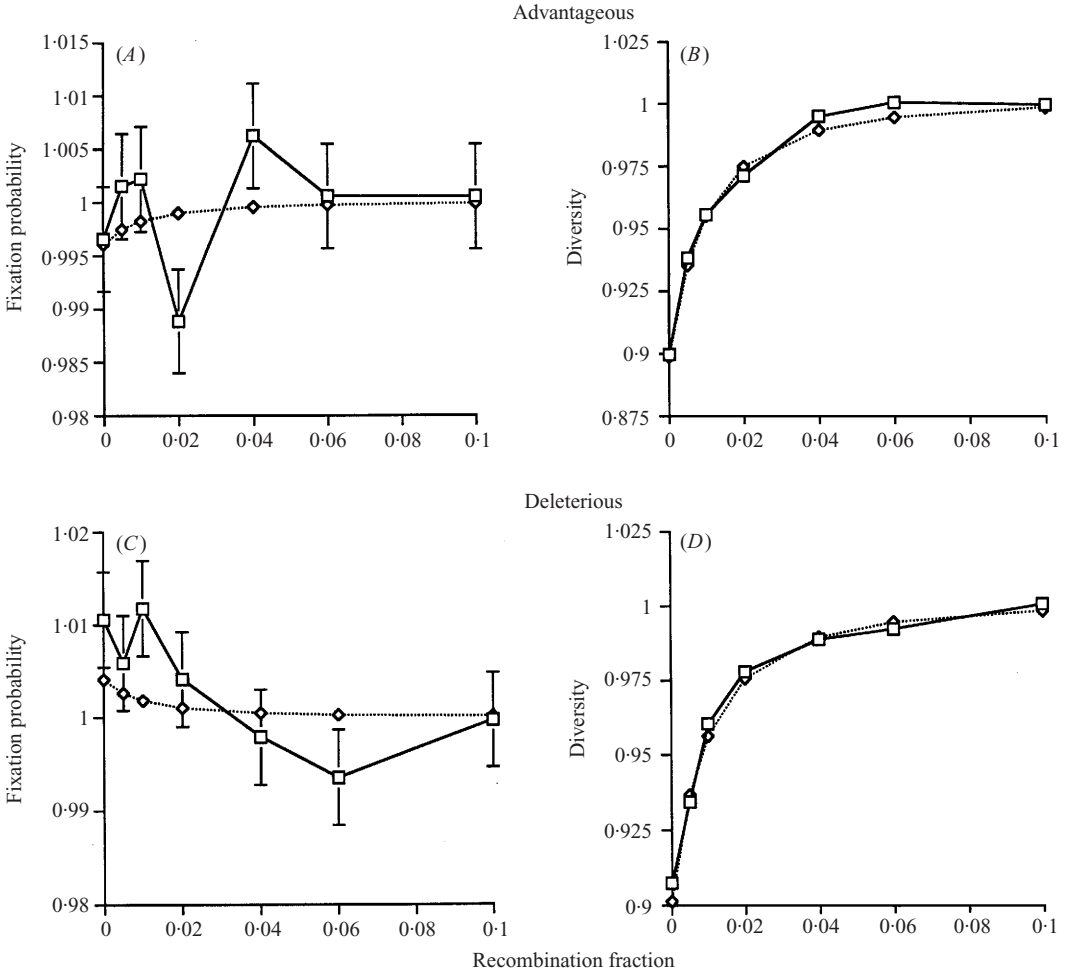

Fig. 1. The effect of background selection on fixation probability (A and C) and nucleotide site diversity (B and D) of advantageous (A and B) and deleterious (C and D) variants with $|Ns| = 0{\cdot}02$. The x-axis indicates the recombination fraction between the locus experiencing recurrent deleterious mutation and the locus of interest. In each figure, simulation results are given by continuous lines (with standard errors) and the values predicted by theory are given by dotted lines. For the cases shown, $N = 2000$, $q = 0{\cdot}1$, $u = 10^{-3}$, $t = 10^{-2}$ and each point represents the average of $1{\cdot}6 \times 10^8$ replications.

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

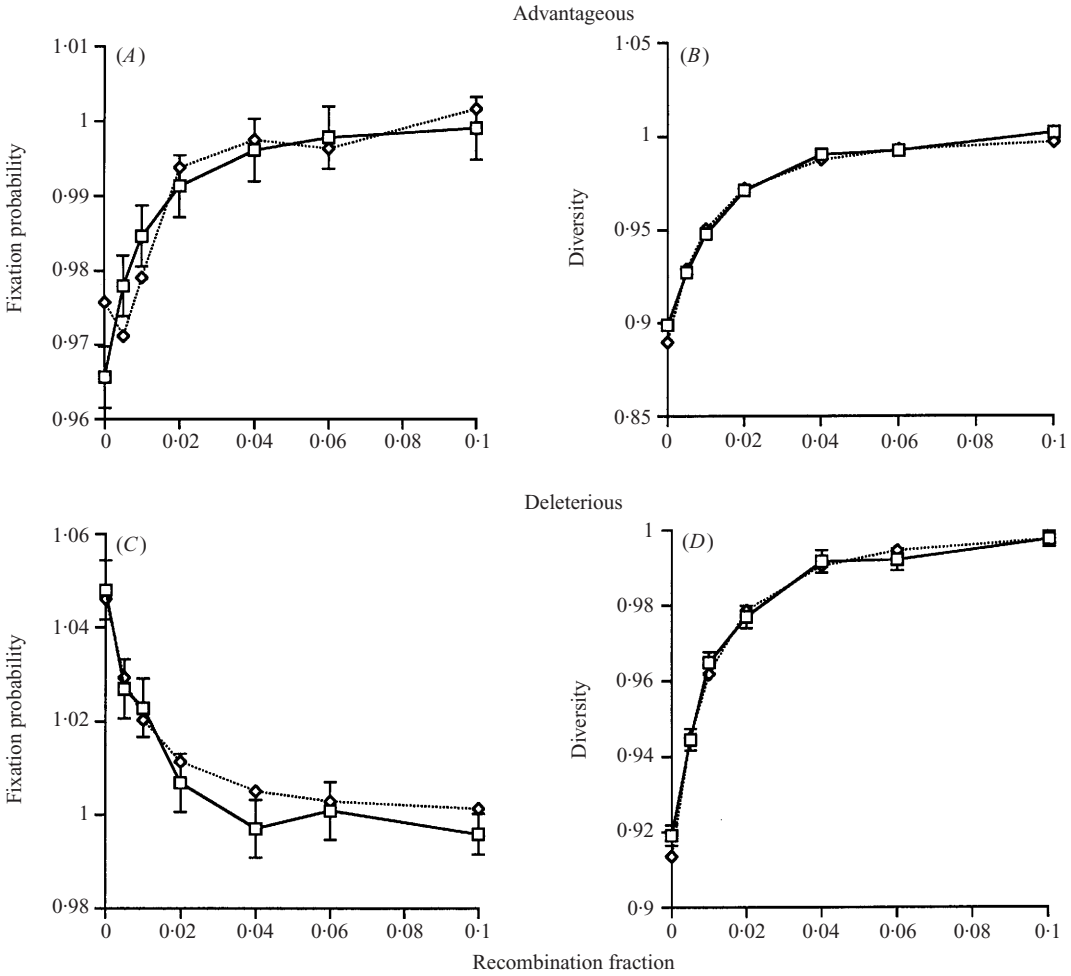

Fig. 2. The effect of background selection on fixation probability (*A* and *C*) and nucleotide site diversity (*B* and *D*) of advantageous (*A* and *B*) and deleterious (*C* and *D*) variants with $|Ns| = 0{\cdot}2$. The other variables are as in Fig. 1.

of the effects on *a* chromosomes (although a variant may move between classes by recombination). The sampling process for each is performed separately, assuming population sizes of Np and Nq, respectively.

In a generation when the number of the rarer allele at the *B* locus in either class with respect to the *A* locus was less than 15, a random number was drawn from a Poisson distribution whose mean was equal to the number of copies of the allele in the parental generation, after the deterministic forces had changed allele frequencies according to (1). This was used to generate the number of alleles of the rarer type (ancestral or mutant) in the next generation. To increase the speed of the simulations, the Ito pseudo-sampling method was employed when the number of alleles of each type was greater than 15 (Li, 1980; Charlesworth *et al.*, 1995). In the Ito method, the change in frequency of an allele is approximated by a stochastic difference equation

$$\Delta x = M_{\delta x} \pm \sqrt{V_{\delta x}}, \tag{27}$$

which has a deterministic component $M_{\delta x}$ due to selection and a stochastic element (due to drift) represented as a random walk of magnitude $\sqrt{V_{\delta x}}$ (where $V_{\delta x}$ is the variance in change in gene frequency due to drift). Uniform random numbers are drawn to decide the sign of the change in gene frequency at the *B* locus, such that negative and positive increments occur with equal frequency. For the *A* class, we have $V_{\delta x} = x(1-x)/2Np$, and for the *a* class, $V_{\delta y} = y(1-y)/2Nq$.

The frequencies of loss and fixation, the times to loss and fixation, and the sum of the heterozygosities for each generation over the sample path for a variant were followed for many replicate introductions (at least 8×10^7). These statistics were then compared with values for mutations with the same selection coefficient when there is no background selection, and with the values predicted by the analytical methods presented in this paper.

At the suggestion of Dr John Gillespie, multi-dimensional diffusion simulations were also performed to investigate the effects of stochastic variation in frequency at the locus experiencing recurrent deleterious mutation. This was achieved by modelling the system under mutation–selection–drift equilibrium for

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

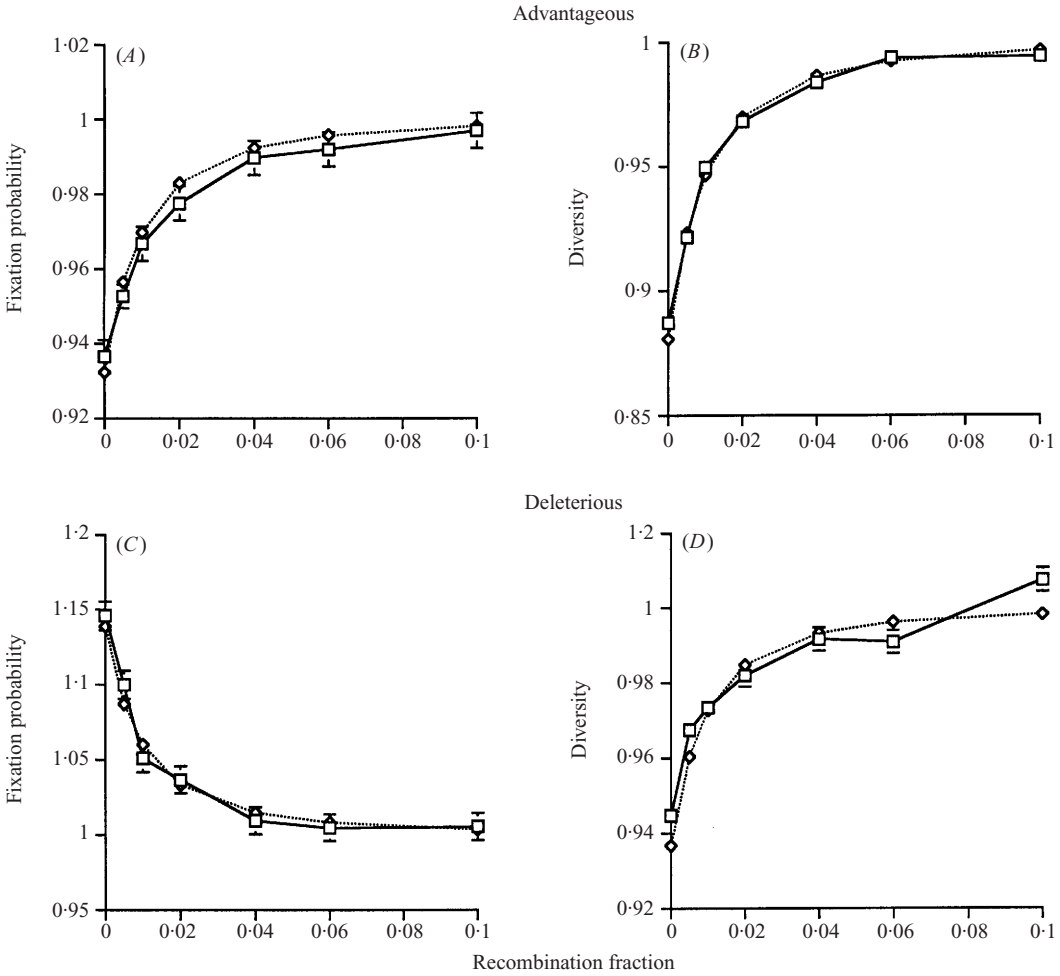

Fig. 3. The effect of background selection on fixation probability (*A* and *C*) and nucleotide site diversity (*B* and *D*) of advantageous (*A* and *B*) and deleterious (*C* and *D*) variants with $|Ns| = 0{\cdot}5$. Each point represents the average of 8×10^8 (advantageous) or $1{\cdot}6 \times 10^8$ (deleterious) replications. The other variables are as in Fig. 1.

both loci simultaneously and calculating the expected heterozygosity and frequency of the preferred allele at the weakly selected locus in a sample from the population every $10N$ generations, after an initial period of $10N$ generations to reach mutation–selection–drift equilibrium. Samples separated by $10N$ generations should be largely independent (Li, 1987). Each generation, the frequencies of the four alleles *AB*, *Ab*, *aB* and *ab* were altered first by selection (according to the formulae in Hill & Robertson, 1966), then by mutation, and finally drift. When every allele was present in at least 20 copies, the Ito pseudo-sampling method was used to simulate the change in frequency due to drift (for the implementation of the Ito scheme in a multidimensional situation see Li, 1980). When any one allele was present in fewer than 20 copies, multinomial sampling was used to obtain the number of copies of the three least frequent allele types.

Simulations were carried out for different values of the selection coefficient at both the weakly selected locus and the locus experiencing recurrent deleterious mutation; dominance was intermediate ($h = 1/2$) for both loci. The mutation rate at the *A* locus was adjusted to give an expected frequency of the deleterious allele, *a*, at mutation–selection balance of 0·1 (using $q = u/t$), except when the selection coefficient was zero (at which point the expected frequency is 1). A much lower reverse mutation rate (*a* to *A*) was also included (10^{-6}) to prevent permanent fixation of the *a* allele when it is under weak selection. The rates of forward and reverse mutation at the weakly selected locus (*B*) were both 10^{-5}, and a population size of $N = 1000$ was used. Only cases of complete linkage between the two loci were considered.

(ii) *Simulation results*

The results for both negative and positive selection on newly introduced alleles at the *B* locus are displayed in Figs. 1–3, for various magnitudes of *Ns*. Nucleotide site diversities and fixation probabilities are plotted

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

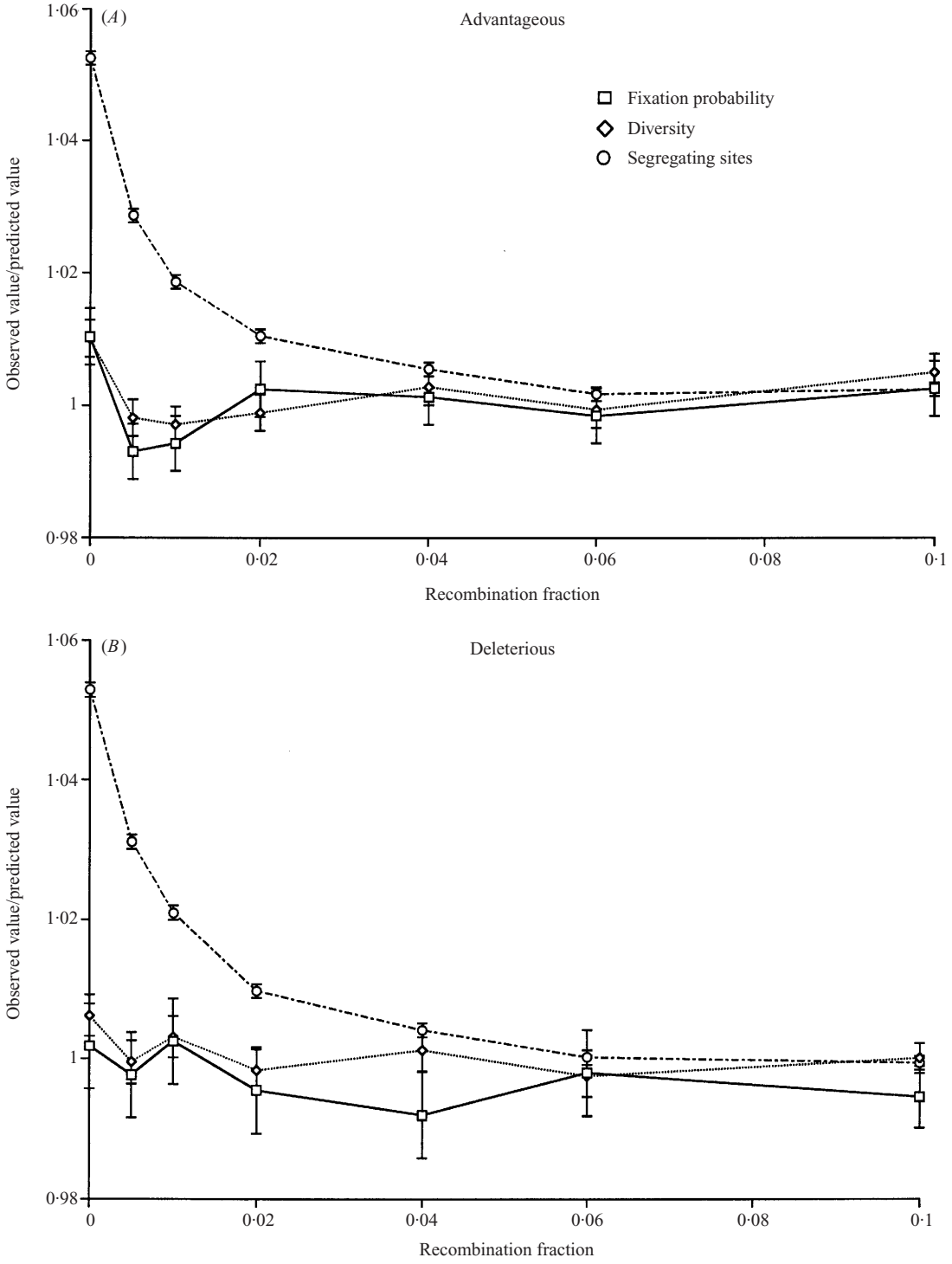

Fig. 4. The discrepancy between the simulation results and the values predicted for the effect of background selection (as modelled by a change in N_e), for fixation probability, nucleotide site diversity and the number of segregating sites in the population for advantageous (*A*) and deleterious (*B*) variants with $|Ns| = 0{\cdot}2$. Parameter values are the same as in Fig. 2. The expected number of segregating sites was calculated from the expected sojourn time in each frequency class, given the reduction in N_e, due to background selection. These values were corrected for a small difference (1%) in the average time to loss or fixation between simulations and numerical estimates in the absence of background selection.

against the recombination fraction r', as the ratio of the observed values to those expected under the same selection model with no background selection. The theoretical values and simulation results are displayed on each plot. Fig. 4 displays results for the fixation probability, nucleotide site diversity and number of segregating sites on the same graph.

The agreement between theory and simulation for

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

Table 1. *Summary of simulation results to examine the effects of background selection on completely linked, weakly selected variants, when there is stochastic variation in allele frequencies at both loci*

Ns at weakly selected site	Nt at linked site	$N\mu$ at linked site	Average frequency preferred allele (F)[a]	Expected F/F_0	F/F_0[b]	Average expected heterozygosity	Expected H/H_0	H/H_0
0	0	2	0·502 ± 0·007		—	0·03796 ± 0·00012		—
	2	0·2	0·500 ± 0·004		0·998 ± 0·016	0·03571 ± 0·00037		0·939 ± 0·012
	10	1	0·491 ± 0·005		0·980 ± 0·017	0·03441 ± 0·00016		0·905 ± 0·008
	20	2	0·507 ± 0·007	1·000	1·011 ± 0·020	0·03456 ± 0·00008	0·900	0·904 ± 0·010
0·2	0	2	0·679 ± 0·003			0·03673 ± 0·00027		—
	2	0·2	0·670 ± 0·004		0·987 ± 0·007	0·03380 ± 0·00030		0·920 ± 0·011
	10	1	0·656 ± 0·003		0·966 ± 0·006	0·03324 ± 0·00025		0·905 ± 0·009
	20	2	0·656 ± 0·005	0·975	0·967 ± 0·009	0·03322 ± 0·00038	0·909	0·904 ± 0·012
0·5	0	2	0·858 ± 0·003			0·03000 ± 0·00031		—
	2	0·2	0·843 ± 0·002		0·982 ± 0·004	0·02866 ± 0·00030		0·955 ± 0·014
	10	1	0·839 ± 0·003		0·977 ± 0·005	0·02784 ± 0·00023		0·928 ± 0·012
	20	2	0·842 ± 0·003	0·974	0·981 ± 0·005	0·02842 ± 0·00026	0·940	0·947 ± 0·013

[a] Each figure is the average across 10 rounds of simulation, each of which consists of 10^4 samples, one taken every $10N$ generations. F is the average frequency of the preferred allele, H is average expected heterozygosity. F_0 and H_0 are the values observed in the absence of background selection. Expected values are calculated from diffusion theory assuming the infinite sites model.

[b] Standard errors calculated by the delta technique.

diversities and fixation probabilities is generally excellent for all values of s and r, except for the probabilities of fixation when the selection coefficient is very small. This is probably due to the increased influence of stochastic forces when the selection coefficient is low (in this case for $|Ns| < 0{\cdot}1$). Given that there is no consistent deviation between the simulated and analytical results, the observed discrepancies are unlikely to be due to the approximation procedure of the analytical results. The agreement between the predicted and observed diversities is excellent for all values of s. In sum, the effects of background selection on these aspects of the behaviour of weakly selected, linked variants can be accurately approximated as a change in N_e, as in the case of linked neutral variants (Charlesworth *et al.*, 1993), and for completely linked, weakly selected variants (Charlesworth, 1994). As expected from the argument in Section 4(ii), the total sojourn times are not well predicted by substituting N_e into the standard diffusion equation formulae, as was previously found in the neutral case (Charlesworth *et al.*, 1993) (Fig. 4).

The results of the simulations which modelled mutation–selection–drift equilibrium at both loci are shown in Table 1. For $|Nt| \geqslant 1$, the average frequency of the deleterious allele at the A locus is in close agreement with the value expected from the deterministic formula $q = u/t$ (data not shown). Furthermore, the observed variance in frequency of the deleterious allele at the A locus is very close to that expected from diffusion theory for $|Nt| \geqslant 1$. It is not then surprising to find that the effect of background selection on the average heterozygosity at the linked, weakly selected locus (B) is in agreement with the results shown previously, when $|Nt| \geqslant 10$ at the A locus. That is, for the case of no recombination, the effect of background selection on weakly selected linked variants can be accurately modelled as a change in N_e. We expect that this should also apply for recombination rates greater than zero. The approximation does not hold only when $|Nt| = 2$ at the A locus.

The effect of background selection on the average frequency of the preferred allele is also consistent with the predictions of the above theory. For $|Ns| = 0$ at the weakly selected locus, background selection has no effect on the average frequency (always 0·5). For $|Ns| > 0$ at locus B, there is a reduction in the average frequency of the preferred allele when $|Nt| > 0$ at the A locus. This is in close agreement with the value expected from the reversible mutation model of selection and drift, using the predicted value of N_e (Li, 1987; Bulmer, 1991). It is worth noting that the predicted effect of a change in N_e on heterozygosity is of a greater relative magnitude than on the average frequency of the favoured allele.

6. Discussion

The general conclusion from the results we have presented is that background selection caused by a single locus affects nucleotide site diversity and fixation probabilities at linked sites subject to weak selection as though effective population size is reduced by the factor f of (23*b*). The same factor has been shown to apply to nucleotide site diversity at neutral sites

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

(Hudson & Kaplan, 1994, 1995; Nordborg *et al*., 1996), and to the probability of fixation of a relatively strongly selected favourable mutation (Barton, 1995). Santiago & Caballero (1998) have pointed out, however, that the formula for f is inexact for large r as far as neutral diversity is concerned (the reduction below 1 is underestimated by a factor of 2 for free recombination). The discrepancy is, however, very hard to detect in the case of background selection at a single locus, where the reduction in N_e is extremely small for large recombination fractions. For this reason, we have confined our simulations to recombination fractions of 10% or less.

The following heuristic argument suggests a reason for the similarity between the neutral and weakly selected cases. In general, under a diffusion model, the effect of drift on the transition from one generation to the next at a locus can be described completely by the sampling variance of allele frequencies at this locus, in the generation under consideration. As shown by Ethier & Nagylaki (1980), weak selection of intensity s at the locus perturbs the sampling variance from the neutral value by a term of order s relative to its value in the absence of such selection. Such a locus should thus behave very similarly to a neutral one as far as the sampling variance due to drift in any one generation is concerned.

It is known that the effects of selection at linked loci on the sampling variance experienced by alleles at a given locus in an arbitrary generation can be described by a sum of cumulative terms involving the effects on this variance of the variance in fitness caused by randomly generated associations with genotypes at the selected locus or loci, over 1, 2, 3… generations previously (Robertson, 1961; Santiago & Caballero, 1998). In the case of background selection due to a single locus, the sum of these terms for a neutral locus converges rapidly, to yield an asymptotic value for the factor by which N_e is reduced. This is identical to f in (23*b*), provided that r is not too large. As shown by Nordborg *et al*. (1996) and Santiago & Caballero (1998), this asymptotic value is sufficient to provide an accurate approximation for the neutral nucleotide site diversity, which is primarily determined by variants that have persisted for some time in the population. The above argument on the effect of weak selection on the sampling variance implies that this result should also hold true as a good approximation for loci subject to weak selection; i.e. their nucleotide site diversity is determined by the standard formula for a weakly selected locus (Kimura, 1983, p. 45), replacing N by Nf.

It is not immediately obvious why this reduction in N_e should also apply to the probability of fixation of a new weakly selected mutant allele, since the fate of a rare variant is strongly affected by stochastic events occurring in the first few generations after its origin. But, as noted by Kimura (1983, p. 229), there is a negligible effect of selection on a rare allele on its distributional properties under drift, since the relevant term is of the order of the product of selection coefficient s and the allele frequency (see the first terms in (1*a*) and (1*b*)). Thus, the effect of weak selection at the B locus on the probability of fixation of a variant at this locus will not become manifest for many generations after its origin by mutation. Combining this result with the argument used above implies that the asymptotic value of N_e derived for the neutral case should apply as a good approximation for the fixation probability of a weakly selected allele. It is remarkable that the branching process analysis of Barton (1995) for the case when $Ns \gg 1$ also yields a similar result (see (23*c*)).

If these heuristic arguments are correct, then it should be possible to extrapolate the conclusions on fixation probabilities and nucleotide diversities at weakly selected sites to multiple loci contributing to background selection, at least if their fitness effects combine multiplicatively, by an extension of the arguments of Nordborg *et al*. (1996) and Santiago & Caballero (1998). It should thus be feasible to extend analyses of the effects of regional differences in recombination rate on nucleotide diversity at neutral sites in the *D. melanogaster* genome (Hudson & Kaplan, 1995; Charlesworth, 1996) to weakly selected sites. In addition, it should also be possible to develop quantitative predictions of the effect of background selection on weak selection for preferred codons, by including sites subject to both favourable and deleterious variants in the same model (cf. Li, 1987; Bulmer, 1991). This will enable us to explore the extent to which the observed relation between recombination and codon bias in *D. melanogaster* (Kliman & Hey, 1993) can be explained by background selection.

One caveat should be noted, however. The results discussed here assume that simultaneous weak selection at multiple sites in the genome has no effects on variation and evolution at individual sites, and that departure from single-locus expectations is caused solely by strong selection at loci linked to the sites in question. But if linkage among weakly selected sites is tight, the Hill–Robertson effect of mutual interference among linked loci under selection (Hill & Robertson, 1966) will also reduce the efficacy of selection (Li, 1987). The likely magnitude of such effects requires further study.

As discussed in Section 4(ii), and confirmed by the simulation results displayed in Fig. 4, we do not expect as large an effect of background selection on the number of segregating sites in the population as on the statistics already described. Santiago & Caballero (1998) have described a method for numerically determining the effect of background

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

selection on the number of segregating sites for the neutral case. This requires the use of a formula for the probability that a neutral variant remains segregating an arbitrary number of generations after its introduction into the population (Voronka & Keller, 1975) as well as expressions for the effects of background selection on the effective population sizes at these times. It is possible to use Kimura's (1955) solution of the forward diffusion equation with genic selection to obtain results for the probability of segregation of a weakly selected variant, but we have not pursued this approach. This is because the number of segregating sites in a sample from a population is much more closely related to the diversity per site than to the number of segregating sites in the population, since rare variants have a low probability of inclusion in a sample. It is hard, therefore, to detect any effect of background selection on the shape of the allele frequency distribution in samples of reasonable size (Hudson & Kaplan, 1995; Charlesworth *et al*., 1995).

It should be noted that the order in which the deterministic and stochastic elements of the Ito method were carried out had an important effect on the fit between theory and simulation. The results shown are for the case where the deterministic portion is evaluated first, which gives excellent agreement with the theory. This also holds for the purely neutral case. When the stochastic element is evaluated first, there is strong disagreement between theory and simulation, particularly at high recombination fractions (where the effect of background selection should be weakest). Biologically, the procedure of carrying out the deterministic changes first is more realistic, since it assumes that the next generation of adults (with a fixed finite population size) is formed by sampling from an infinite pool of gametes, whose composition is controlled by the effects of the deterministic forces on the genotype of the adults of the previous generation (Ethier & Nagylaki, 1980; Nagylaki, 1990).

Inspection of (1) suggest that the reason for this discrepancy is the possibility of a significant effect of the terms in $x-y$ on the allele frequencies in the initial generations after introduction of a variant, before recombination has reduced the value of $x-y$. The terms in question are likely to be small relative to the stochastic changes in x or y if r is small (since q and t are assumed to be small), but may be comparable to the stochastic changes if r is large. For example, if a B mutation is introduced initially into the a class, $x = 0$ and $y = 1/(2Nq)$ in the first generation. Hence, the deterministic changes in x and y are dominated by $r/(2N)$ and $p(r+t)/(2Nq)$, respectively. If random sampling is applied after these deterministic changes, the corresponding stochastic changes have standard deviations of approximately $r^{0.5}/(2N)$ and $1/(2Nq)$. These measure the expected magnitude of the stochastic changes. If random sampling is applied before the deterministic changes, the standard deviations are 0 and (approximately) $1/(2Nq)$, respectively. There is thus a significantly larger probable stochastic effect on x in the first case. A similar argument applies to the stochastic effects on y when a B mutation is introduced initially into the A class.

This raises the question of why the diffusion equation approximations seem to work so well, since formally they require the expected effects of the deterministic forces to be sufficiently small that second-order terms in them and their product with the sampling variance in allele frequencies can be neglected in comparison with first-order terms (Ewens, 1979, chap. 4). Inspection of (1) indicates that, while this condition can easily be met for the frequency x of B among A chromosomes by making q sufficiently small, this is not necessarily true for the frequency y of B among a chromosomes. If r and t are sufficiently small, this condition can be met for (1*b*) (assuming that s is small), but it is not necessarily met if either of these is large, unless $|x-y|$ is small compared with y. For recombination fractions of more than a few per cent, the diffusion approximation is thus likely to be inaccurate in the initial generations, before $|x-y|$ is reduced to a small value by recombination. For this reason, caution should be exercised in using (22) for situations in which r is more than a few per cent. The practical effect of this is likely to be small, however, since the effect of background selection with t values of realistic magnitude decreased rapidly as r increases (Figs. 1–3).

Appendix

(i) *Derivation of equation* (*18*)

We note that $e^{L_1\tau}(y-y(x))^n - p_x(y)$ is the solution of the equation

$$\frac{\partial}{\partial\tau}f(y,\tau) = Lf(y,\tau) \tag{A 1}$$

with the initial condition

$$f(y,0) = (y-y(x))^n p_x(y). \tag{A 2}$$

Hence,

$$e^{L_1\tau}(y-y(x))^n p_x(y) = \int_0^1 dy'\, p(y,\tau\,|\,y',0) \times (y'-y(x))^n p_x(y'), \tag{A 3}$$

and

$$D_n(x) = \int_0^\infty d\tau \int_0^1 dy'\, dy (y-y(x))\, p(y,\tau\,|\,y',0) \times (y'-y(x))^n p_x(y'). \tag{A 4}$$

https://doi.org/10.1017/S0016672399003705 Published online by Cambridge University Press

The integral over y produces the trajectory of the first moment. We obtain an appropriate expression by integrating the ordinary differential equations corresponding to (1). These can be written as

$$\frac{d}{d\tau}(y-y(x)) = -(y-y(x)) - \frac{s}{b}x(1-x) + \frac{s}{b}y(1-y). \tag{A 5}$$

Linearizing $y(1-y)$ about $y = y(x)$, taking the expectation, and using the formulae (A 7) and (A 8) for the first and second stationary moments from Section (ii) of the Appendix, we find for the time-dependent first moment near the quasi-equilibrium $y = y(x)$

$$\frac{d}{d\tau}E(y-y(x) = -\left(1-\frac{s}{b}(1-2y(x))\right) \times E(y-y(x)) + O\left(\left(\frac{s}{b}\right)^2, \frac{s}{b}\frac{1}{\beta}\right). \tag{A 6}$$

Integrating this ODE with the initial condition $y'-y(x)$, and carrying out the remaining integrations in (A 4) leads to (18).

(ii) *The stationary moments*

The stationary moments can be computed by linearizing the exponential function in (14) (because $\alpha \ll 1$) and by repeatedly exploiting the well-known property of the gamma function, $\Gamma(z+1) = z\Gamma(z)$. The resulting formulae are relatively simple, since $\beta \gg 1$ (or, equivalently, $2Nu \gg 1$) may be assumed for most biological applications (Nordborg *et al*., 1996). The lowest-order terms of the stationary moments are

$$E_{ss}(y) = y(x) = x + \frac{s}{b}x(1-x), \tag{A 7}$$

$$E_{ss}((y-y(x))^2) = \frac{1}{2\beta}x(1-x), \tag{A 8}$$

$$E_{ss}((y-y(x))^3) = \frac{1}{2\beta^2}x(1-x)(1-2x), \tag{A 9}$$

and

$$E_{ss}((y-y(x))^4) = \frac{3}{4\beta^2}x^2(1-x)^2. \tag{A 10}$$

This work was supported by grants from the Royal Society to B. C. and the National Environment Research Council to B. C. and Deborah Charlesworth, and by NSF grants DEB-9407226 and DEB-9896179 to W. S. We thank John Gillespie and an anonymous reviewer for suggesting improvements to the manuscript.

References

Aguadé, M. & Langley, C. H. (1994). Polymorphism and divergence in regions of low recombination in *Drosophila*. In *Non-neutral Evolution: Theories and Molecular Data* (ed. B. Golding), pp. 67–76. London: Chapman and Hall.

Akashi, H. (1995). Inferring weak selection from patterns of polymorphism and divergence at 'silent' sites in *Drosophila* DNA. *Genetics* **139**, 1067–1076.

Akashi, H. (1996). Molecular evolution between *Drosophila melanogaster* and *D. simulans*; reduced codon bias, faster rates of amino-acid substitution, and larger proteins in *D. melanogaster*. *Genetics* **144**, 1297–1307.

Akashi, H. & Schaeffer, S. (1997). Natural selection and the frequency distributions of 'silent' DNA polymorphism in *Drosophila*. *Genetics* **146**, 289–307.

Aquadro, C. F., Begun, D. J. & Kindahl, E. C. (1994). Selection, recombination, and DNA polymorphism in *Drosophila*. In *Non-neutral Evolution: Theories and Molecular Data* (ed. B. Golding), pp. 46–56. London: Chapman and Hall.

Barton, N. H. (1995). Linkage and the limits to natural selection. *Genetics* **140**, 821–884.

Bulmer, M. G. (1991). The selection–mutation–drift theory of synonymous codon usage. *Genetics* **129**, 897–207.

Charlesworth, B. (1994). The effect of background selection against deleterious alleles on weakly selected, linked variants. *Genetical Research* **63**, 213–228.

Charlesworth, B. (1996). Background selection and patterns of genetic diversity in *Drosophila melanogaster*. *Genetical Research* **68**, 131–150.

Charlesworth, B., Morgan, M. T. & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303.

Charlesworth, D., Charlesworth, B. & Morgan, M. T. (1995). The pattern of neutral molecular variation under the background selection model. *Genetics* **141**, 1619–1632.

Dvorak, J., Luo, M.-C. & Yang, Z.-L. (1998). Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing *Aegilops* species. *Genetics* **148**, 423–434.

Ethier, S. & Nagylaki, T. (1980). Diffusion approximations of Markov chains with two time scales and applications to population genetics. *Advances in Applied Probability* **12**, 14–19.

Ewens, W. J. (1979). *Mathematical Population Genetics*. Berlin: Springer-Verlag.

Gardiner, C. W. (1990). *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Berlin: Springer-Verlag.

Gillespie, J. H. (1994). Alternatives to the neutral theory. In *Non-neutral Evolution: Theories and Molecular Data* (ed. B. Golding), pp. 1–17. London: Chapman and Hall.

Gillespie, J. H. (1997). Junk ain't what junk does: neutral alleles in a selected context. *Gene* **205**, 291–299.

Hill, W. G. & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research* **8**, 269–294.

Hudson, R. R. (1994). How can the low levels of DNA sequence variation in regions of the *Drosophila* genome with low recombination rates be explained? *Proceedings of the National Academy of Sciences of the USA* **91**, 6815–6818.

Hudson, R. R. & Kaplan, N. L. (1994). Gene trees with background selection. In *Non-neutral Evolution: Theories and Molecular Data* (ed. B. Golding), pp. 140–153. London: Chapman and Hall.

Hudson, R. R. & Kaplan, N. L. (1995). Deleterious background selection with recombination. *Genetics* **141**, 1605–1617.

Kaplan, N. L., Hudson, R. R. & Langley, C. H. (1989). The 'hitch-hiking' effect revisited. *Genetics* **123**, 887–899.

Kimura, M. (1955). Stochastic processes and distribution of gene frequencies under natural selection. *Cold Spring Harbor Symposia on Quantitative Biology* **20**, 33–53.
Kimura, M. (1964). *Diffusion Models in Population Genetics*. London: Methuen.
Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press.
Kimura, M. (1985). The role of compensatory neutral mutations in molecular evolution. *Journal of Genetics* **64**, 7–19.
Kliman, R. M. & Hey, J. (1993). Reduced natural selection associated with low recombination in *Drosophila melanogaster*. *Molecular Biology and Evolution* **10**, 1239–1258.
Li, W.-H. (1980). Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. *Genetics* **95**, 237–258.
Li, W.-H. (1987). Models of nearly neutral mutations with particular implications for non-random usage of synonymous codons. *Journal of Molecular Evolution* **24**, 337–345.
Maynard Smith, J. & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetical Research* **23**, 23–35.
Moriyama, E. N. & Powell, J. R. (1996). Intraspecific nuclear DNA variation in *Drosophila*. *Molecular Biology and Evolution* **13**, 261–277.
Nachman, M. W. (1997). Patterns of DNA variability at X-linked loci in *Mus domesticus*. *Genetics* **147**, 1303–1318.
Nagylaki, T. (1990). Models and approximations for random genetic drift. *Theoretical Population Biology* **37**, 192–212.
Nordborg, M., Charlesworth, B. & Charlesworth, D. (1996). The effect of recombination on background selection. *Genetical Research* **67**, 159–174.
Peck, J. (1994). A ruby in the rubbish: beneficial mutations, deleterious mutations, and the evolution of sex. *Genetics* **137**, 597–606.
Robertson, A. (1961). Inbreeding in artificial selection programmes. *Genetical Research* **2**, 189–194.
Santiago, E. & Caballero, A. (1998). Effective size and polymorphism of linked neutral loci in populations under selection. *Genetics* **149**, 2105–2117.
Stephan, W. (1995). An improved method for estimating the rate of fixation of favorable mutations based on DNA polymorphism data. *Molecular Biology and Evolution* **12**, 959–962.
Stephan, W. (1996). The rate of compensatory evolution. *Genetics* **144**, 419–426.
Stephan, W. & Langley, C. H. (1998). DNA polymorphism in *Lycopersicon* and crossing-over per physical length. *Genetics* **150**, 1585–1593.
Stephan, W., Wiehe, T. H. E. & Lenz, M. W. (1992). The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. *Theoretical Population Biology* **41**, 237–254.
Voronka, R. & Keller, J. B. (1975). Asymptotic analysis of stochastic models in population genetics. *Mathematical Biosciences* **25**, 331–362.