J. theor. Biol. (1995) **173,** 307–320

Periodic Selection and Hitchhiking in a Bacterial Population

Otto G. Berg

Department of Molecular Biology, University of Uppsala Biomedical Center, Box 590, *S*-75124 *Uppsala, Sweden*

(*Received on* 19 *May* 1994, *Accepted on* 19 *October* 1994)

The accumulation of a neutral marker in a bacterial population under balanced growth in a chemostat follows a jagged curve as adaptive variants continuously appear and sweep the population. Such periodic selection curves are simulated in the present work using deterministic equations that, in contrast to previous models, take full account of the stochastic character of the process. The uncertainties due to the appearance time, the survival probability, and the extent of early growth at small numbers are included as stochastic initial conditions for every new variant—adaptive or neutral—that appears. The model is used to calculate the substitution rate via hitchhiking where a neutral or weakly selected mutation is carried along when a new adaptive one takes over the population. The expected ratio for the probabilities of the presence or absence of a weakly selected or counterselected mutation in the population is also calculated. This can be related to the standard result without hitchhiking if the average time between adaptive shifts is interpreted as an effective population size.

1. Introduction

The neutral theory of evolution (Kimura, 1983) holds that most genetic changes at the molecular level are nonadaptive and their ultimate fixation depends on random drift. Because of their short generation times, bacterial cultures offer one way of testing some tenets of the neutral theory. However, due to their often very large sizes (10^8–10^{10} individuals), the population dynamics of a bacterial culture behaves qualitatively differently in many respects from the much smaller populations usually considered in this context.

In a chemostat a bacterial culture can be kept continuously in balanced exponential growth. The growth rate constant is determined by the dilution rate so that the total number of cells in the culture is approximately constant (for reviews, see e.g. Kubitschek, 1970; Dykhuizen & Hartl, 1983). The dynamics of the population changes under these conditions can be investigated for instance by studying the mutational appearance and accumulation of a neutral marker in the cell population. Experimentally, it is found that the accumulation of the neutral marker follows a jagged curve with periods of steady rise followed by rapid falls. This phenomenon has been called *periodic selection* and is caused by the regular mutational appearance and subsequent exponential take-over of the population by new variants with a selective advantage over the original cells. Since a new adaptive variant is most likely to occur in the large fraction of cells without neutral marker, the take-over (*adaptive shift*) of the population leads to a quick decrease in the fraction of cells with the neutral marker. When the new variant has become established, the neutral marker will start to accumulate by mutation again. In chemostat cultures that have been followed for very long times, there seems to be little or no decrease in the rates of occurrence of these adaptive shifts (Helling *et al*., 1987).

In a very large population, recurrent mutations would lead to a steady accumulation of a neutral variant. If adaptive shifts do not occur, the fraction of cells carrying a neutral marker would, after very long time, approach a level where its accumulation is balanced by the mutational loss. Without adaptive shifts, a neutral variant would never (within astronomical time) be able to take over a very large

0022–5193/95/070307+14 $08.00/0

population through random drift only. The adaptive shifts have two consequences. First, the mutational accumulation of cells carrying the neutral marker will never reach large values; second, in the rare cases when the adaptive mutation occurs in the small fraction carrying a neutral mutation, this can become dominant in the population by "hitchhiking" with the adaptive one (Maynard Smith & Haigh, 1974; Kaplan *et al*., 1989). Similar interrupted growth behavior can be expected for weakly selected or counterselected mutant variants.

This behavior of neutral mutations stands in contrast with some formulations of the theory of evolution (Kimura, 1983), where single mutations are assumed to be fixed via genetic drift without interference from other mutations. This difference in point of view occurs because the neutral theory has been developed mostly for small populations or for sexual populations where the fate of a certain mutation is largely unlinked to that of others. The hitchhiking possibility allows neutral or weakly counterselected mutations to become fixed also in a large population where genetic drift by itself (i.e. uninfluenced by other mutations) would take a very long time. However, it is not obvious what the actual rates of fixation will be in such a case. Ultimately, the fixation probabilities will determine the patterns of genetic variability in the surviving cells.

In order to calculate the hitchhiking probabilities under periodic selection we need a consistent statistical description of the population dynamics of the contributing genetic variants. Traditionally, one would employ deterministic growth equations to describe large populations; this is satisfactory as long as the statistical variations around the average numbers are small. However, in an adaptive event, the population of the new variant starts off with such small numbers that a stochastic description is required. The stochastic equations are notoriously much more difficult to handle, particularly if many variants are involved, and a single event in periodic selection involves at least four different genetic variants.

The main focus of this paper is to develop a stochastic theory that can describe the dynamics of a large asexual population where mutations are strongly linked and the fixation of one necessarily influences all others. For completeness, we shall first look at the well-known deterministic growth equations and discuss how they fail. Then we shall consider the stochastic growth equations and derive some useful results from them. This will be done mostly in the Appendix, where a formalism is developed to account for the statistical effects due to timing in appearance and probability of survival. To describe the periodic selection, these two approaches will be combined in a model where the deterministic growth is "stitched" together with a stochastic appearance of the new variants. The predictions of the model can be tested against the experimental periodic selection curves and used to interpret their shapes. Then the model is used as a basis to calculate the hitchhiking probability and fixation rates for weakly selected or counterselected mutations.

2. Deterministic Growth

Consider a population with a constant total size of N_T cells which at some initial time $t=0$ is dominated by one variant, (0). At later times mutations appear and accumulate: variant (1) appears with mutation rate constant u_1 and has a growth rate advantage s_1 relative to variant (0), variant (2) appears with rate u_2 in the (1)-population and has growth rate advantage s_2 relative to that, etc. for variants (3) and higher. In this succession of strongly selected variants weakly selected or counterselected (possibly neutral) mutations also appear. We will be interested in a particular one: $(0)\rightarrow(0w)$, $(1)\rightarrow(1w)$, etc. which appears with mutation rate constant u_w and has a growth rate advantage s_w relative to its parent. Schematically, the process can be described as:

$$\begin{array}{ccccccc} (0) & \xrightarrow{u_1} & (1) & \xrightarrow{u_2} & (2) & \xrightarrow{u_3} & (3) \xrightarrow{u_4} \\ \downarrow u_w & & \downarrow u_w & & \downarrow u_w & & \downarrow u_w \\ (0w) & & (1w) & & (2w) & & (3w) \end{array} \qquad (1)$$

In principle, there are back mutations as well as adaptive mutations also in the variants carrying the (w)-mutation. However, as long as the total fraction of this variant remains small, these will be very rare events. Below we shall therefore consider these rare events only in the context of hitchhiking.

Introducing the specific growth rate constants k_i, k_{iw} for variants (i), (iw), and the constant dilution rate D for the whole system, the following equations for the numbers $N_i(t)$, $N_{iw}(t)$ of the different variant are obtained:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t}=(k_i-D)N_i+u_iN_{i-1} \qquad (2a)$$

$$\frac{\mathrm{d}N_{iw}}{\mathrm{d}t}=(k_{iw}-D)N_{iw}+u_wN_i. \qquad (2b)$$

The solution of these equations is formally simple; the only problem is that the growth rate constants k_i and k_{iw} must depend on the composition of

the population. It is therefore more convenient to consider the ratios

$$g_i(t) = N_i(t)/N_{i-1}(t) \quad \text{and} \quad f_i(t) = N_{iw}(t)/N_i(t) \tag{3}$$

The resulting differential equations for these ratios are

$$\frac{dg_i}{dt} = s_i g_i + u_i - u_{i-1}\frac{g_i}{g_{i-1}} \tag{4a}$$

$$\frac{df_i}{dt} = s_w f_i + u_w - u_i \frac{f_i}{g_i}, \tag{4b}$$

where the growth rate advantages

$$s_i = k_i - k_{i-1}, \quad s_w = k_{iw} - k_i \tag{4c}$$

are assumed to be constants. In principle, these equations can be solved recursively for an arbitrary initial condition. From the definitions, eqn (3) above, one finds that the overall fraction of cells that carry the weakly selected or counterselected mutation (w) is given by

$$F_w(t) = \frac{f_0 + g_1\{f_1 + g_2[f_2 + g_3(f_3 + g_4(f_4 + \cdots))]\}}{1 + f_0 + g_1\{1 + f_1 + g_2[1 + f_2 + g_3(1 + f_3 + g_4(1 + f_4 + \cdots))]\}}. \tag{5}$$

This will display a steady growth also if one considers the more complete growth equations that include the back mutations etc. Thus, the deterministic equations do not generate the characteristic periodic selection pattern observed in experiments for a neutral marker. This is because they allow exponential growth also of such small fractions of the population that correspond to much less than a single cell. This unrealistic description does not cause any problems when the rate of mutational appearance is so large that it takes a very short time to get the first individuals present. However, in the case considered here, new variants start at very small numbers and the variants that derive from them must have a very large uncertainty in their time of appearance. Exponential growth and accumulation cannot start until the first surviving mutant has appeared. The deterministic equations in fact predict that the later variants become established and extinct again long before the probability of their first appearance has become significant.

In the following it will be assumed that the rate of mutational appearance of the selected variants is so small that it is sufficient to consider only one event. This is not strictly true in general, but, as is shown in the Appendix, the statistical behavior is so strongly dominated by the appearance of the first mutant that the description of the continued growth is not much affected by this assumption. We can thus neglect the terms involving u_i in the differential equations (4) and they can be easily solved:

$$g_i(t) = g_i^0 \exp(s_i t) \tag{6a}$$

$$f_i(t) = \exp(s_w t)\left(f_i^0 + \frac{u_w}{s_w}[1 - \exp(-s_w t)]\right) \approx f_i^0 + u_w t. \tag{6b}$$

The approximation in the last equation is valid for a neutral mutation where, by definition, s_w is vanishingly small. After its establishment, each new selected variant, (i), grows exponentially until it takes over the population, or until the next selected variant, ($i+1$), starts to contribute. A neutral variant that appears in the (i)-population will first experience an exponential inflation phase, if it occurs before variant (i) has taken over, while the continued growth will be linear in time after the take-over by variant (i), as indicated by eqn (6b).

3. Stochastic Growth

The stochastic description is concerned with the probability distribution of the numbers of the different variants at different times in the population. A single event in a series of periodic selections requires one to consider at least four different variants: the original one and the selected one plus the neutral marker in both. During the early growth phase, the uncertainty in the number of cells of a certain variant is usually much larger than the expected average. This is due mostly to the timing of the first appearance, to the large probability of ultimate loss, and to the random nature of the early growth of small numbers. At later times after the variant has grown to a sufficiently large number of cells, the continued growth can be described with the deterministic equations since only small variations around the average values will occur. The stochastic effects then enter primarily as initial conditions in the deterministic results. The basic theory required to describe these effects has been developed in the Appendix.

A birth-and-death process is used to calculate the probability distribution for the number of descendants of a new variant that has appeared in some background of a changing composition of previous variants. Since the growth and survival of the new variant depends on the genetic composition of the population, this will

strongly depend on its time of appearance. Most new mutants will be diluted out without leaving any descendents at later times; these will have little or no influence on the genetic composition of the population. Thus it is convenient to consider the distribution of descendants *conditional on survival*; the others will not contribute to the changing composition of the population. The time of appearance of a new variant that leaves surviving descendants can be determined from the mutation rate constant, the size of the population of the parental variant, and the survival probability. This gives a probability distribution for the appearance time. The probability distribution for the number of descendants, conditional on survival, can also be calculated. This gives both the average amount of growth and the statistical variation around it. The average growth is the same as that given by the deterministic equations, if the initial condition is chosen appropriately.

Consider the appearance and growth of the new selected variant (i) after the time when variant ($i-1$) first appeared. The distribution in the time of appearance for the first surviving mutant (i.e. one that leaves surviving descendents after some length of time) of variant (i) is calculated in the Appendix. This distribution is plotted in Fig. 1 for some reasonable choices of parameters. When $u_i N_T/D$ increases the distribution becomes sharper, and for smaller values of $u_i N_T/D$ the statistical variance in the size of the time delay increases. Although mutations can appear at early times before the distribution in Fig. 1 becomes appreciable, such events are very rare and extremely unlikely to appear in a periodic selection experiment. However, their possible occurrence is included in the simulations by using the formula

$$T_i = \frac{1}{s_{i-1}} \ln\left[(rnd)^{-s_{i-1}(1+D/s_i)/u_i N_T} \times \left(1 + \frac{1}{g^0_{i-1}} \frac{1+D/(s_{i-1}+s_i)}{1+D/s_i}\right) - \frac{1}{g^0_{i-1}} \frac{1+D/(s_{i-1}+s_i)}{1+D/s_i} \right] \quad (7)$$

to generate the stochastic time delay for the appearance of variant (i). The parameter g^0_{i-1} is the initial condition of the parental variant ($i-1$), and the function *rnd* is the output of a random number generator between 0 and 1. As shown in Fig. 1, a large number of delay times generated by eqn (7) conforms to the distribution calculated according to eqn (A.21) of the Appendix. (See, for example, Fluendy, 1970, on how to generate the stochastic numbers from the distribution function.)

There is also a large stochastic component to the amount of growth during early stages when the number of cells is small. Since the probability distribution for the actual number of surviving descendants is found to be approximately exponential [eqn. (A.11) in the Appendix] this uncertainty can be accounted for simply by multiplying a stochastic factor:

$$S = -\ln(rnd) \quad (8)$$

to the calculated average number of cells. This procedure gives, with an exponential probability distribution, the actual number of cells present in one particular realization of the growth kinetics.

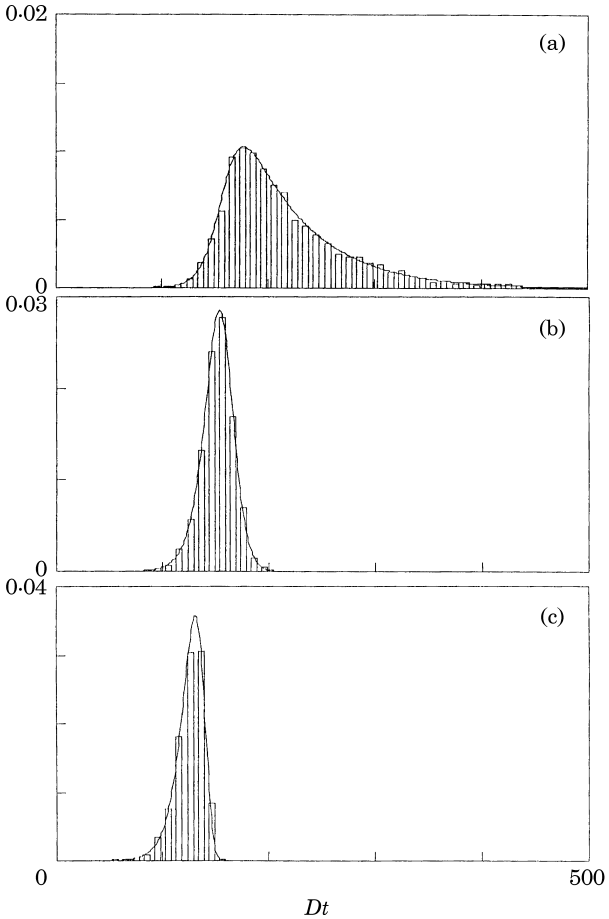

FIG. 1. Distribution of appearance times for a new selected variant. The histograms are based on 5000 events generated with eqn (7) and the smooth curves are the expected distributions according to eqn (A.21). Data used in the three case are $s_{i-1}/D = 0.1$, $s_i/D = 0.2$, and $g^0_{i-1} = 10^{-7}$. Panel (a) is for $u_i N_T/D = 0.1$, (b) for $u_i N_T/D = 1$, and (c) for $u_i N_T/D = 10$. The numbers on the time axis correspond to the time in units of $1/D$; the time in units of average generation time would be this number divided by ln(2).

Under periodic selection, these adaptive shifts where new selected variants replace their predecessors are studied with the help of a neutral marker (variant (w) with $s_w = 0$). The stochastic time delay before the neutral marker appears in the $(i-1)$ population can be calculated from the probability given in eqn (A.37) of the Appendix:

$$T_{i-1,w} = \frac{1}{s_{i-1}} \ln\left[1 - \frac{D + s_{i-1}}{u_w N_T g_{i-1}^0} \ln(rnd)\right]. \qquad (9)$$

This expresses the time delay of the first surviving neutral variant after variant $(i-1)$ has become established. There is also a large stochastic component in the amount of early growth, which can be accounted for by the same factor as in eqn (8).

After the first surviving mutant of variant (i) has appeared at time T_i the probability distribution for the number of descendants at any later time can be calculated as described in the Appendix. Consider some time later when the average number of descendants is fairly large (say 10^4) but still a small fraction of the total. From this time on, the statistical variance around the average is small and the continued growth can be considered as deterministic. However, the actual number of cells at this time is not the average but some stochastic number distributed around the average. Thus the initial condition is the expected average multiplied by the stochastic factor given in eqn (8). The deterministic equations can also be extrapolated backwards to the time of appearance, and the whole growth can be described by the deterministic equations with a stochastic starting time and a stochastic initial condition. It should be noted that in this way, the initial condition for each variant at the appearance time is not simply one cell; the expectation value from the stochastic equations conform to the deterministic growth equations only with an initial condition that depends on, for example, the appearance time and growth advantages. This description does not account in detail for what actually happens during the early growth of a new variant, but the consequences for the later growth are appropriately included.

4. Deterministic Growth with Stochastic Delays

A periodic selection experiment can be started by first allowing the population of cells to grow up to the density that the chemostat will carry at the given dilution rate. Assume that the initial clone is of type (0). During the growth some may mutate to another selected variant, (1). If it assumed that this first phase is unlimited exponential growth with growth rate constant k_0, the time to reach the required population size will be about $(1/k_0)\ln(N_T)$. During this time, a variant (1) is likely to occur at the latest after time $(1/k_0)\ln(k_0/u_1)$. Variant (1) will grow with rate $k_1 = k_0 + s_1$. Thus at time $t = 0$ when the experiment starts, there are mostly cells of type (0) but also some fraction of type (1):

$$\frac{N_1(0)}{N_0(0)} = g_1^0 \approx \frac{u_1}{k_0}\left(\frac{u_1 N_T}{k_0}\right)^{s_1}. \qquad (10)$$

Thus if $u_1 N_T$ is greater than k_0, the next selected variant has with a large probability appeared already before time $t = 0$. Similarly, the neutral variant under consideration will also most likely become established during the set-up stage. Thus for the appearance of variants 1 and $0w$ there will be no time delays. After time $t = 0$ the growth is described by the balanced exponential with constant dilution rate, D, in accordance with eqn (6) above. One finds:

$$N_0(t) = N_T \frac{1}{1 + g_1[1 + g_2(1 + g_3[1 + g_4(1 + \cdots)])]} \qquad (11a)$$

$$N_i(t) = g_i(t) N_{i-1}(t), \quad \text{for } i = 1, 2, \ldots \qquad (11b)$$

$$N_{0w} = u_w t N_0(t) \qquad (12a)$$

$$N_{1w} = [f_1^0 + u_w(t - T_{1w})] N_1(t) H(t - T_{1w}) \qquad (12b)$$

$$N_{iw} = [f_1^0 + u_w(t - t_i^{\text{app}} - T_{iw})] N_i(t) \times H(t - t_i^{\text{app}} - T_{iw}), \quad \text{for } i = 2, 3, \ldots \qquad (12c)$$

where

$$g_1(t) = g_1^0 \exp(s_1 t) \qquad (13a)$$

$$g_i(t) = g_i^0 \exp[s_i(t - t_i^{\text{app}})] H(t - t_i^{\text{app}}), \quad \text{for } i = 2, 3, \ldots \qquad (13b)$$

The appearance time of variant (i) is determined by the appearance time for variant $(i-1)$ plus the stochastic time delay from eqn (7):

$$t_i^{\text{app}} = t_{i-1}^{\text{app}} + T_i \quad \text{for } i = 2, 3, \ldots \qquad (13c)$$

These relations assume that the neutral marker does not reach a significant fraction of the total population, so that, to a good approximation, the total size is normalized only by the constraint $N_0(t) + N_1(t) + N_2(t) + \cdots = N_T$. (This assumption holds for the periodic selection experiments; it can easily be relaxed if necessary.) The Heaviside function, $H(t - t_i)$, which is equal to 1 when $t > t_i$ and 0 otherwise, has been introduced as a factor to emphasize the fact that the new variants cannot contribute significantly to the population before they

have become established. The stochastic time delays for the different selected variants are determined from eqn (7) above: the time delay is counted from the time that the previous selected variant became established. The extra stochastic time delays for the neutral marker are calculated from eqn (9). The initial condition for g_1, which is probably established before time $t=0$, is given by eqn (10). The initial conditions for the later selected variants ($i=2, 3,\ldots$) are determined by eqn (A.24) of the Appendix, while the inital conditions for the neutral marker are given by eqn (A.35). These initial conditions are to be multiplied with the stochastic factor, eqn (8), to account for the large uncertainty during the early growth. In this way one can continue to build up the successive variants from the properties of the preceding ones. The fraction of the total population that carries the neutral marker (w) is given by eqn (5).

Figure 2 shows one simulation of the expected patterns in a series of take-over events based on eqns (11–13). The three phases of growth of the neutral variant are clearly seen: first exponential inflation together with the parent, then steady accumulation after the parent has taken over, and finally exponential decline when the new selected variant appears. The parameter values used in Fig. 2 have been chosen to generate curves similar to the ones from the longest running experiment with *E. coli* (Helling *et al.*, 1987). Some of the adaptive shifts merge into each other and are not visible as independent events. Thus, the observed number of peaks may underestimate the number of adaptive shifts; this is the case particularly for large values of $u_i N_T/D$. The stochastic time delays, both for the selected and the neutral variants, are crucial in order to generate curves similar to the experimental ones. With $u_i N_T/D$ and $u_w N_T/D$ large, the effective time delays become more predictable and less varied, but the uncertainty in the amount of early growth [eqn (8)] remains. If $u_i N_T/D$ and $u_w N_T/D$ are small, independent experiments on the same system would give drastically different curves. In the simulations it sometimes happens (about once per 50–100 adaptive shifts, data not shown) that the uncertainty produces such a large peak for the neutral marker that it dwarfs the other peaks.

The strong stochastic character of the curves makes it difficult, if not impossible, to interpret the shape of an individual take-over event in terms of the underlying growth parameters. The most obvious relationship is the slope of the rising parts that is determined by u_w. However, some other useful average relations can also be found. Two consecutive adaptive shifts are shown in Fig. 3. The times indicated on the abscissa are the various stochastic delay times

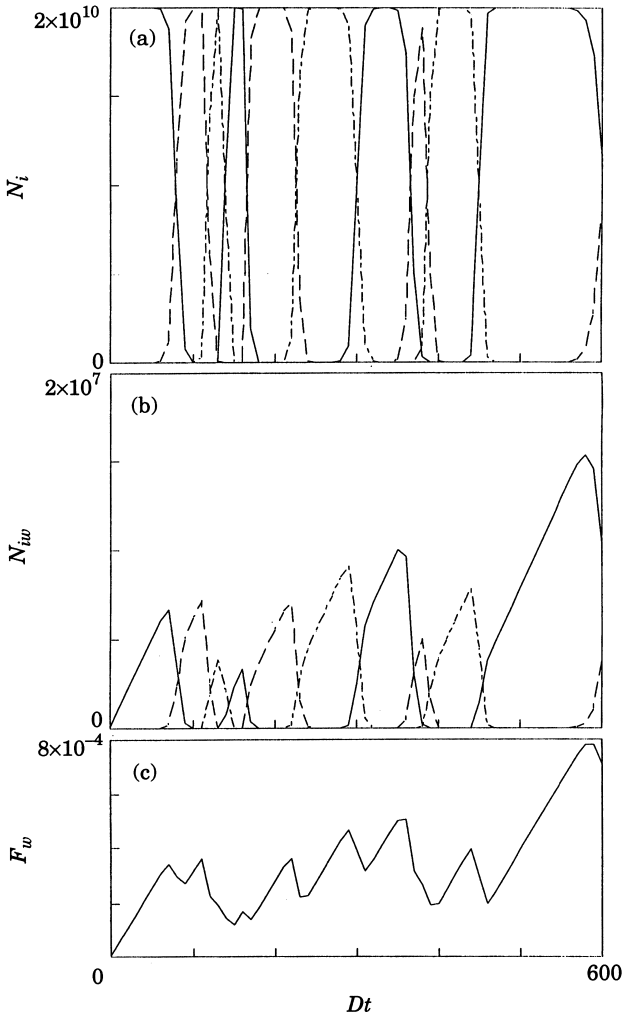

FIG. 2. Simulated periodic selection curves. Panel (a) shows the growth and decline of the selected variants (11 variants, $i=0, 1, 2,\ldots, 10$ shown). Panel (b) shows the growth and decline of the neutral variant derived from these selected ones. Panel (c) shows the fraction of the population that carries the neutral marker. Data used in the simulation: $N_T=2\ 10^{10}$; $u_i/D=5\ 10^{-11}$ (all i); $u_w/D=5\ 10^{-6}$; $s_i/D=0.3,\ 0.5,\ 1.0,\ 0.8,\ 0.4,\ 0.3,\ 0.3,\ 0.4,\ 0.15$ (for $i=1, 2,\ldots, 10$). The numbers on the time axis correspond to the time in units of $1/D$.

described above. The extrapolation to the abscissa of the linear part of the growth curve for the neutral marker gives approximately its stochastic appearance time, at least on the average. The most useful parameter that can be found graphically in the figure is the time between the appearance of the neutral marker and the time that its parent reaches dominance. As discussed in the caption to Fig. 3, this is

$$\Delta T_i=\frac{1}{s_i}\left[1+\ln\left(\frac{u_w N_T}{D+s_i}\right)\right], \tag{14}$$

which can be used to find an estimate for s_i, since u_w is known from the slope and N_T can be calculated from

the cell density. Kubitschek (1974) has carried out a similar analysis, but without taking the time delay of the neutral variant into account. Thereby the factor u_w/D is missing in his corresponding equation which consequently would give a very much larger estimate for the growth advantage s_i. Equation (14) is only the most likely value; the stochastic uncertainties in the appearance time and early growth of the neutral variant still remain. Another complication is the occurrence of incomplete take-overs where a selected variant is overtaken by its successor before it has become dominant; use of eqn (14) in such cases can also distort the estimate of the growth advantage.

In one study of periodic selection in yeast (Paquin & Adams, 1983*a*) the growth advantages of successive adaptive mutants were measured to be around $s/D=0.1$. However, the experimental curves have very small ΔT, which using eqn (14) would suggest that $s/D=0.8$ or more. The source of this discrepancy is not known but could perhaps be explained if the growth advantage, s, for each selected variant decreases as the amount of that variant increases in the population. In this way, fixation and early growth could be determined by a larger value of s than that measured in competition experiments.

The general upwards trend over many adaptive shifts that is often observed for the presence of a neutral marker in *E. coli* can be explained by a decrease in the selective value for successive adaptive variants (cf. Kubitschek, 1974). The average time between adaptive shifts is most sensitive to the growth rate advantage, s, of the new variants. If both s and the total population size has been determined, the rate of appearance of adaptive variants can be estimated from the frequency of the adaptive shifts according to eqn (A.26–A.27) of the Appendix.

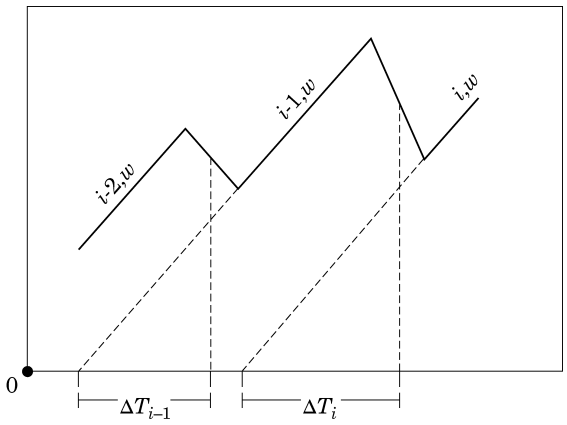

FIG. 3. Frequency of neutral marker through two consequtive adaptive shifts. The rising parts of the curves are for variants $(i-2, w)$, $(i-1, w)$, and (i, w) as indicated; these have the slope equal to u_w/D. Time zero has arbitrarily been set at the time when variant $i-1$ is established. The falling parts of the curves are when the new selected variants $i-1$ and i, respectively, take over; the take-over for variant $i-1$ at the half-way fall occurs at time $(1/s_{i-1})\ln(1/g^0_{i-1})$. The rising part of variant $(i-1, w)$ extrapolates to the x-axis at time [cf. eqns (A.36), (A.39) and (A.40)] $T_{i-1,w}-1/s_{i-1}\approx(1/s_{i-1})[\ln((D+s_{i-1})/u_wN_Tg^0_{i-1})-1]$. Thus the time interval between these events is $\Delta T_{i-1}=(1/s_{i-1})[1+\ln(u_wN_T/(D+s_{i-1}))]$. Similarly, the corresponding time interval ΔT_i for variant i is as given by eqn (14).

5. Hitchhiking

In the following we will consider the more general situation where (w) is a weakly selected or counterselected mutation. After variant $(i-1)$ has become established, the (w)-mutation, $(i-1, w)$, will start to accumulate in this subpopulation. On average, the fraction of the $(i-1)$-population that carries (w) is given by f_{i-1} from eqn (3). If the next selected variant (i) occurs at time t' after the establishment of variant (i), the probability that it will occur in a cell that carries (w) is given on the average by:

$$f_{i-1}(t')=\frac{u_w}{s_w}\,[\exp(s_wt')-1]. \tag{15}$$

In this equation it is appropriate to use the overall average f_{i-1}—rather than its stochastic representation—since the system will go through a large number of adaptive shifts before a hitchhiking event actually takes place. If the adaptive shift occurs on the average at time $\bar{T}$ after the establishment of the previous variant, the hitchhiking probability for variant (w) will be

$$P_H(s_w)=\frac{u_w}{s_w}\,[\exp(s_w\bar{T})-1]. \tag{16}$$

This is valid as long as $P_H\ll 1$; otherwise (w) corresponds to a strongly selected variant that would be fixed by itself rather than via hitchhiking. In this way, the hitchhiking probability of eqn (16) refers to an average over all shifts and not just to variant (i). If u_iN_T is small so that effectively only one clone contributes to the new population, this hitchhiking probability will also express the probability that the (w)-variant becomes dominant in the adaptive shift.

It is interesting to consider the ratio of the hitchhiking probabilities for variants with a weak selective advantage s_w and a weak selective disadvantage $-s_w$. In a population where (w) is largely absent, a (w)-mutant has advantage s_w, while in a population dominated by (w), a mutational loss of (w) would have disadvantage $-s_w$. This will give the long time average for the ratio of the probabilities that the weakly selected mutation is present or absent. Using eqn (16)

one finds the approximation:

$$\frac{P_H(s_w)}{P_H(-s_w)} \approx \frac{u_w}{v_w} \exp[s_w \bar{T}], \tag{17}$$

where v_w is the mutation rate constant for the loss of variant (w). $\bar{T}$ is the average time between adaptive shifts. This result can be compared with the standard result (e.g. Bulmer, 1991) for the mutation-selection balance for the ratio of the presence and absence of a weakly selected variant:

$$\frac{P(s_w)}{P(-s_w)} = \frac{u_w}{v_w} \exp\left(\frac{s_w}{D} N_e\right), \tag{18}$$

where N_e is the effective population size (a factor 2 has been incorporated). Since s_w is the growth rate advantage (per unit time), s_w/D corresponds to the relative selection advantage of the (w)-variant over its parent. Thus for a large population with frequent adaptive shifts, the effective population size for this particular application is given by the average time between the adaptive shifts (if $u_i N_T \ll D$, $\bar{T}$ is from eqn (A.26) in the Appendix):

$$\bar{T} \approx \frac{\ln(N_T)}{s} + \frac{s+D}{uN_T s}. \tag{19a}$$

This is valid if the successive selected variants have a similar growth advantage s and adaptive mutation rate constant u, and if uN_T is smaller than D. The first term expresses the time required to penetrate the population (the fixation time) and the second term is the time required before the appearance of the first surviving selected mutation. Thus, a comparison of eqns (17) and (18) shows that the effective population size for the hitchhiking system would be

$$N_e = D\bar{T} \approx \frac{D}{s} \ln(N_T) + \frac{1 + D/s}{uN_T}. \tag{19b}$$

The larger the population is, the more frequent are the adaptive shifts and the smaller is the apparent effective population size. The adaptive shifts and hitchhiking lead to a substantial decrease in the effective population size (Koch, 1974).

If adaptive shifts occur with rate $1/\bar{T}$, and the hitchhiking probability for a weak mutation is given by eqn (16), the effective substitution rate via hitchhiking is approximately

$$K = \frac{P_H(s_w)}{\bar{T}} \approx \frac{u_w}{s_w \bar{T}} [\exp(s_w \bar{T}) - 1]. \tag{20}$$

With $/s_w/\bar{T} \ll 1$, this gives $K = u_w$ as expected for a neutral mutation (Kimura, 1983), and by definition this limit is considered a nearly neutral mutant. Equation (20) is not valid in the limit $s_w \bar{T} \gg 1$, which corresponds to a strongly selected variant and is more likely to be fixed by itself rather than via hitchhiking. Without hitchhiking, the expected substitution rate would be

$$K = \frac{uN_e s_w/D}{1 - \exp(-s_w N_e/D)}. \tag{21}$$

A comparison between eqns (20) and (21) shows that the rate of fixation via hitchhiking for a strongly counterselected variant ($s_w < 0$, $/s_w/\bar{T} > 1$) will be much faster and less sensitive to the value of s_w than is the case without it. While the steady-state distribution, eqns (17) and (18), has exactly the same dependence on s_w (if N_e is identified as $\bar{T}$), the effective substitution rates depend quite differently on s_w in the two cases with and without hitchhiking.

In many unicellular organisms, codon bias seems to be under weak selection pressure and synonymous substitutions would therefore be a good candidate for a process that is driven by hitchhiking. However, in a study of the synonymous substitution rates in *E. coli* and *S. typhimurium* (O. G. Berg & M. Martelius, in preparation), the uncertainties in the data resolution does not yet allow a clear distinction between the results with and without hitchhiking.

6. Discussion and Conclusions

Earlier descriptions of the periodic selection curves by Koch (1974) and Kubitschek (1974) also rely on deterministic equations with time delays for the appearance of the selected variants. However, their calculations did not include any estimates of the size of the expected time delays. As a consequence, much larger growth advantages for the selected variants were required in order to generate curves that resemble the experimental ones. Levin (1981) has also simulated a run of replacements by higher-fitness clones. However, his emphasis was not on describing the stochastic effects and generating periodic selection curves, but on studying the relationship between hitchhiking, recombination, and genetic population structure.

The present model provides a consistent description of the strong stochastic effects involved in the appearance and early growth of new variants in a population that is undergoing a change in composition. The growth curves can very easily be simulated on a computer since all relations have been determined explicitly and no numerical integrations are required: apart from the exponential growth equations given above, only a random number generator is needed to determine the appearance times and initial conditions for the new variants. One complication in modeling the adaptive shifts is that the

growth rate constants must change in response to the changing composition of the population. This can easily be handled for the deterministic growth equations simply be studying the ratios of the different variants. This approach is not very useful for the stochastic growth equations. Instead a formalism is developed in the Appendix that allows the calculation of the stochastic effects in the appearance and early growth of new variants, also when the growth rates are time dependent. The main difference between the model presented here and the previous ones, is the explicit use of the stochastic growth equations which provide the required probability distributions. This makes it very simple to account properly for the uncertainty in the time of appearance as well as in the amount of early growth for all new variants. Unfortunately, the large statistical variances expected preclude a detailed interpretation of the selection parameters based on the shape of the individual events in the periodic selection curves.

The deterministic equations with stochastic time delays give rise to the characteristic periodic selection curves. Although the equations allow very short time delays, these are extremely rare events and do not show up in the simulations or in the experiments. However, if they do occur (*jackpots*), their number of descendants would also be very large, so their influence on the average is still considerable. This is why the proper statistical average of the stochastic equations agrees with the prediction from the deterministic equations, and still does not give a useful description of the adaptive shifts.

The assumptions of the model imply that each new adaptive variant grows faster by a certain amount than its parent. Obviously, this improvement cannot continue indefinitely. However, since each new variant grows only in the background of its immediate one or two predecessors, it is only compared to them that it is necessarily better. It could in fact be growing slower in a competition with some of its more distant predecessors as has been observed (Paquin & Adams, 1983*b*). Thus there is no one optimal phenotype under these circumstances: the fitness of a variant is determined by its competition. This allows the possibility that for any set of circumstances (existing variants and growth conditions), there may always exist better variants that can appear through mutation, and periodic selection could be cycling through adaptive shifts indefinitely without actually achieving a long-term improvement in fitness (Paquin & Adams, 1983*b*).

The hitchhiking relations allow an identification of the effective population size with the average time between adaptive shifts. Equation (18) has been applied in a model describing codon bias (Bulmer, 1991) and in a model for the bias-pair variability in gene-regulatory binding sites (Berg, 1992). In both cases it was found that the effective population size would be around 10^4–10^5 for *E. coli*, which is in many orders of magnitude smaller than the number usually expected (Ochman & Wilson, 1987). In the light of the present results, this relatively small number could be an indication that natural populations of *E. coli* undergo adaptive shifts every 10^4–10^5 generations—these would indeed be rare events.

In eqn (19b) we have used the expected average from the periodic selection in a constant environment where adaptive variants appear through mutation and stochastic growth. Natural populations may experience adaptive shifts also through changes in the environment or through other processes that forces the population through severe "bottle necks". Thus, the average time between adaptive shifts that determines the effective population size of a natural population would have to include all processes whereby a new variant can take over and bring with it a random sampling of the weakly selected or counterselected variations that has accumulated in a large population. The effective population size that appears in these relations would refer to the effective population size of individual lineages. A statistical comparison of the differences between lineages may well involve some other measure of the effective population size (cf. Levin, 1981).

I thank Julian Adams, Måns Ehrenberg and Chuck Kurland for comments on an earlier version of the manuscript. This work was supported by the Swedish Natural Science Research Council.

REFERENCES

Berg, O. G. (1992). The evolutionary selection of DNA base pairs in gene-regulatory binding sites. *Proc. natn. Acad. Sci. U.S.A.* **89,** 7501–7505.

Bulmer, M. (1991). The selection-mutation drift theory of synonymous codon usage. *Genetics* **129,** 897–907.

Dykhuizen, D. E. & Hartl, D. L. (1983). Selection in chemostats. *Microbiol. Rev.* **47,** 150–168.

Fluendy, M. (1970). Monte Carlo methods. In: *Markov Chains and Monte Carlo Calculations in Polymer Science* (Lowry, G. G., ed.) pp. 46–90. New York: Marcel Dekker.

Harris, T. E. (1963). *The Theory of Branching Processes*. Berlin: Springer.

Helling, R. B., Vargas, C. N. & Adams, J. (1987). Evolution of *Escherichia coli* during growth in a constant environment. *Genetics* **116,** 349–358.

Kaplan, N. L., Hudson, R. R. & Langley, C. H. (1989). The "hitchhiking" effect revisited. *Genetics* **123,** 887–889.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge Universtiy Press.

Koch, A. L. (1974). The pertinence of the periodic selection phenomenon to prokaryotic evolution. *Genetics* **77,** 127–142.

KUBITSCHEK, H. E. (1970). *Introduction to Research with Continuous Cultures*. Englewood Cliffs, NJ: Prentice-Hall.
KUBITSCHEK, H. E. (1974). Operation of selection pressure on microbial populations. *Symp. Soc. Gen. Microbiol.* **24,** 105–130.
LEVIN, B. R. (1981). Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. *Genetics* **99,** 1–23.
MAYNARD SMITH, J. & HAIGH, J. (1974). The hitch-hiking effect of a favourable gene. *Genet. Res.* **23,** 23–35.
OCHMAN, H. & WILSON, A. C. (1987). Evolutionary history of enteric bacteria. In: *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C., ed.) pp. 1649–1654. Washington, DC: ASM Press.
PAQUIN, C. & ADAMS, J. (1983*a*). Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. *Nature, Lond.* **302,** 495–500.
PAQUIN, C. E. & Adams, J. (1983*b*). Relative fitness can decrease in evolving asexual populations of *S. cerevisiae*. *Nature, Lond.* **306,** 368–371.

APPENDIX

Descendants from a Single Individual

Consider a population of cells where at some initial time a new variant i appears. This has the growth rate at constant k_i. Assume that the population size is kept fixed by a constant dilution rate D. As a consequence, the growth rate constants will adapt to the changing composition of the population so that the average growth equals the dilution. If one individual of variant i appears at time $t=t_i$, then the probability $p_n(t)$ that the population contains n descendants from this mutant at later times is described by the birth and death process:

$$\frac{\mathrm{d}p_n}{\mathrm{d}t}=k_i(n-1)p_{n-1}-k_inp_n - Dnp_n+D(n+1)p_{n+1}. \quad \text{(A.1)}$$

There are several complications with this description. First, there is no cut-off when n approaches the maximal population size N_T. Second, the growth rate, k_i, for the new variant i will depend on the composition of the population in which it grows: ultimately, when n becomes large, k_i should also depend on n. Thus, the equations are adequate only as long as n is much smaller than N_T. Below, we shall use them only to describe the first establishment of a new variant in a large population that is undergoing a change in composition of other variants. In this way, the growth rate constant k_i will, in general, be time dependent. The assumption that the total population size, N_T, is constant is a convenient boundary condition for the deterministic equations: as long as N_T is very large, the assumption is of little consequence for the results.

Introducing the *excess growth rate* relative to the dilution rate at time t

$$\xi_i(t)=k_i(t)-D,$$

one finds that the average number of descendants at later times is

$$\langle n_i(t;\,t_i)\rangle=\exp\left(\int_{t_i}^{t}\xi_i(x)\,\mathrm{d}x\right). \quad \text{(A.2)}$$

The relative variance is

$$\frac{\sigma_i^2(t;\,t_i)}{\langle n_i(t;\,t_i)\rangle^2}=2D\int_{t_i}^{t}\exp\left(-\int_{t_i}^{t'}\xi_i(x)\,\mathrm{d}x\right)\mathrm{d}t' +1-\exp\left(-\int_{t_i}^{t}\xi_i(x)\,\mathrm{d}x\right) \quad \text{(A.3)}$$

The master equations (A.1) can be solved exactly (Harris, 1963) and the result can be written in the form:

$$p_n(t;\,t_i)=\left(1-\frac{q_s(t;\,t_i)}{\langle n_i(t;\,t_i)\rangle}\right)^{n-1}\frac{q_s^2(t;\,t_i)}{\langle n_i(t;\,t_i)\rangle} \quad \text{for } n>0, \quad \text{(A.4)}$$

where

$$q_s(t;\,t_i)=\frac{1}{1+D\displaystyle\int_{t_i}^{t}\exp\left[-\int_{t_i}^{t'}\xi_i(x)\,\mathrm{d}x\right]\mathrm{d}t'} \quad \text{(A.5)}$$

is the *survival probability*, i.e. the probability that the mutant appearing at time t_i has surviving descendents at time t.

If there are n individuals of variant i present at time t, one finds that the probability of survival at later time ($t'>t$) is

$$q_s(t';\,t;\,n)=1-[1-q_s(t';\,t)]^n. \quad \text{(A.6)}$$

The average number of descendants is simply n times the result (A.2), valid for the case of a single cell initially, and the relative variance is $(1/n)$ times the result (A.3). Thus, once the number has reached a certain level, the continued growth can be described deterministically with a small relative statistical variance.

In most cases when a new variant appears, it will quickly disappear from the population leaving no descendants. Only a small fraction, corresponding to $q_s(t;\,t_i)$ for large t, of the mutational appearances will ultimately survive. The transient presence of a small number of mutants that leave no descendants in the long run is not important for the establishment of new variants. Thus it is more convenient to limit the description to the probability, $p_n^{(s)}(t;\,t_i)$, of having n descendants at time t *conditional on ultimate survival*.

If the survival probability after long times is denoted $q_s^\infty(t_i)$, the following equality must hold (Bayes' theorem)

$$p_n^{(s)}(t;t_i)=\frac{q_s^\infty(t;n)p_n(t;t_i)}{q_s^\infty(t_i)}, \tag{A.7}$$

where

$$q_s^\infty(t;n)=1-[1-q_s^\infty(t)]^n \tag{A.8}$$

is the ultimate survival probability conditional on having n present at time t as given by eqn (A.6) above for large t'. Thus, the probability for n descendants at time t conditional on ultimate survival can be written

$$p_n^{(s)}(t;t')=\left[1-\frac{q_s(t;t_i)}{\langle n_i(t;t_i)\rangle}\right]^{n-1} \times\frac{q_s(t;t_i)}{\langle n_i(t;t_i)\rangle}\frac{q_s(t;t_i)}{q_s^\infty(t_i)}[1-(1-q_s^\infty(t))^n]. \tag{A.9}$$

For large n, this is approximately an exponential distribution where the relative variance conditional on survival is nearly equal to 1; thus, the uncertainty in the amount of early growth is large. In the following, we will be interested in the limit of such large t that the average number of descendants is large and the survival probability $q_s(t;t_i)$ approaches its asymptotic value $q_s^\infty(t_i)$. In this limit, the average number of descendants conditional on survival is

$$\langle n_i(t;t_i)\rangle_s=\langle n_i(t;t_i)\rangle/q_s^\infty(t_i). \tag{A.10}$$

If this is the average number, the actual number of descendants could well be quite different according to the distribution (A.9). For large n this distribution can be approximated as an exponential

$$p_n^{(s)}(t;t_i)=\frac{1}{\langle n(t;t_i)\rangle_s}\exp(-(n-1)/\langle n(t;t_i)\rangle_s), \tag{A.11}$$

and the actual number of surviving descendants can be accounted for by multiplying the expected number with the stochastic factor (Fluendy, 1970) given by eqn (8) of the main text. Using this approximation also for small n overestimates their overall probability by approximately $1/\langle n\rangle$ which by assumption is negligibly small.

Continuous Appearance of Mutants

Assume that variant i appears continuously by mutation at a rate of $M_i(t)$ new individuals per unit time. If the probability for n individuals of variant i at time t is $P_n(t)$, then this will satisfy the same master equations as above [eqn (A.1)] after addition of the source terms $M_i(t)P_{n-1}(t)-M_i(t)P_n(t)$:

$$\frac{dP_n}{dt}=k_i(n-1)P_{n-1}-k_inP_n-DnP_n +D(n+1)P_{n+1}+M_i(P_{n-1}-P_n). \tag{A.12}$$

Since variant i is assumed to derive by mutation from variant $i-1$, the rate of mutational appearance at time t' is the mutation rate constant u_i multiplied by the number of cells of variant $i-1$; i.e. $M_i(t')=u_iN_{i-1}(t')$. the average and the variance in the number of mutants present at time t are

$$\langle n_i(t)\rangle=\int_0^t M_i(t')\exp\left[\int_{t'}^t\xi_i(x)\,dx\right]dt' \tag{A.13}$$

$$\sigma_i^2(t)=\int_0^t[(2D+\xi_i(t'))\langle n_i(t')\rangle +M_1(t')]\exp\left[2\int_{t'}^t\xi_i(x)\,dx\right]dt'. \tag{A.14}$$

The probability that no mutant with surviving descendants has appeared before time t is given by

$$Q_0(t)=\exp\left[-\int_0^t M_i(t')q_s(t;t')\,dt'\right], \tag{A.15}$$

where q_s is given by eqn (A.5). If we are interested in the probability that no mutant has appeared by time t that leaves descendents that survive, $q_s(t;t')$ should be replaced by $q_s^\infty(t')$ in the equation above.

Growth of a Selected Variant

Consider a situation where variant $i-1$ has been established at time $t=0$ and is growing deterministically in a background of $i-2$. From eqn (6a) of the main text this gives:

$$N_{i-1}(t)=\frac{N_Tg_{i-1}^0\exp(s_{i-1}t)}{1+g_{i-1}^0\exp(s_{i-1}t)}. \tag{A.16}$$

As a consequence, the excess growth rate for variant $i-1$ is found to be

$$k_{i-1}(t)-D=\xi_{i-1}(t)=\frac{s_{i-1}}{1+g_{i-1}^0\exp(s_{i-1}t)}. \tag{A.17}$$

Sometime during this growth, variant i will appear. As long as its numbers remain small, the excess growth

rate for this variant will be

$$k_i(t)-D=\xi_i(t)=s_i+\frac{s_{i-1}}{1+g^0_{i-1}\exp(s_{i-1}t)} \quad \text{(A.18)}$$

The ultimate survival probability for a mutant of type i appearing at time t_i, as given by eqn (A.5) for large t, is

$$q^\infty_s(t_i)=\frac{g^0_{i-1}+\exp(-s_{i-1}t_i)}{g^0_{i-1}\left[1+\dfrac{D}{s_i}\right]+\exp(-s_{i-1}t_i)\left[1+\dfrac{D}{s_{i-1}+s_i}\right]}. \quad \text{(A.19)}$$

The rate of appearance is $M_i=u_iN_{i-1}(t)$ and the probability that no ultimately surviving i mutant has appeared before time t can be calculated from eqn (A.15):

$$Q^\infty_0(t)=\left[\frac{g^0_{i-1}(1+D/s_i)+1+D/(s_{i-1}+s_i)}{g^0_{i-1}(1+D/s_i)\exp(s_{i-1}t)+1+D/(s_{i-1}+s_i)}\right]^{\frac{s_iu_iN_T}{s_{i-1}(s_i+D)}}. \quad \text{(A.20)}$$

From this, the probability distribution for the time of appearance of the first surviving i-mutant can be calculated as

$$\rho(t)=-\frac{\mathrm{d}Q^\infty_0}{\mathrm{d}t}, \quad \text{(A.21)}$$

which has been plotted in Fig. 1 for some different values of the parameters involved. For a mutant i that appears at time t_i, the average number of descendants at time t conditional on ultimate survival can be calculated using eqns (A.2), (A.9) and (A.19):

$$\langle n_i(t;t_i)\rangle_s=\langle n_i(t;t_i)\rangle/q^\infty_s(t_i)=\exp(s_i(t-t_i))\,\frac{g^0_{i-1}(1+D/s_i)+\exp(-s_{i-1}t_i)(1+D/(s_i+s_{i-1}))}{g^0_{i-1}+\exp(-s_{i-1}t)}. \quad \text{(A.22)}$$

This will follow the same kind of exponential growth as expected from the deterministic equations. The result is valid for the initial growth before variant i has become a significant fraction of the total population. However, since the result is equivalent in form to the deterministic one, at later times one can simply continue with the deterministic growth equations. In terms of the fraction $g_i=N_i/N_{i-1}$ [eqns (3) and (6a) of the main text] this gives at times later than t_i:

$$g_i(t)=g^0_i\exp(s_i(t-t_i)), \quad \text{(A.23)}$$

where the initial condition is given by

$$g^0_i=\frac{1}{N_T}\left[1+\frac{D}{s_i}+\frac{\exp(-s_{i-1}t_i)}{g^0_{i-1}}\left(1+\frac{D}{s_i+s_{i-1}}\right)\right]. \quad \text{(A.24)}$$

Thus for every selective variant i, one can calculate the stochastic time delay T_i from the distribution function equation (A.20). Using this time delay instead of t_i in eqn (A.23–24), the expected deterministic growth follows. The results for each variant i depends on the one immediately preceding it. The growth of the new selected variant is exponential with an initial condition, $g^0_i=g_i(t_i)$, that depends on the time of appearance. This initial condition determines the average growth. Since the actual number of cells present could be quite different from this average according to the distribution in eqn (A.10), the initial condition should be multiplied by a stochastic factor as given by eqn (8) of the main text.

If u_iN_T/D is small, most appearance times t_i will be large and

$$g^0_i\approx\frac{s_i+D}{s_iN_T} \quad \text{for } u_iN_T\ll D. \quad \text{(A.25)}$$

Also in this limit, the distribution of appearance times [eqn (A.21)] will effectively become exponential after a lag time (cf. Fig. 1). The average appearance time will be the sum of the lag time and the time constant of the exponential:

$$\bar{T}_i\approx\frac{1}{s_{i-1}}\ln\left[N_T\frac{1+D/(s_{i-1}+s_i)}{(1+D/s_{i-1})(1+D/s_i)}\right]+\frac{1+D/s_i}{u_iN_T} \quad \text{for } u_iN_T\ll D. \quad \text{(A.26)}$$

If, on the other hand, u_iN_T is large, the distribution of appearance times is strongly peaked around

$$t_i=\bar{T}_i\approx\frac{1}{s_{i-1}}\ln\left[\frac{s_{i-1}(D+s_{i-1}+s_i)}{(s_{i-1}+s_i)u_iN_Gg^0_{i-1}}\right] \quad \text{for } u_iN_T\gg D, \quad \text{(A.27)}$$

and the initial condition is

$$g^0_i\approx\frac{u_i}{s_{i-1}} \quad \text{for } u_iN_T\gg D. \quad \text{(A.28)}$$

The deterministic equations determine the ratios between successive variants: to account for the constant size of the total population, the numbers of all variants present at one time are added and set equal to N_T. In this way the deterministic growth equations

can be normalized, while the stochastic growth equations are used only in the beginning to determine the time of establishment and initial growth before a variant has become a significant fraction of the total population.

These results can be compared to the expected average from the stochastic equations, eqn (A.13), which, using eqn (A.18), gives

$$\langle n_i(t)\rangle = N_{i-1}(t)\,\frac{u_i}{s_i}\,(\exp(s_i t)-1) \qquad \text{(A.29)}$$

With reasonable parameter values, this can be shown to be indistinguishable from the average over the time delays:

$$\langle n_i(t)\rangle = \int_0^t \langle n_i(t;\,t')\rangle_s \rho(t')\,\mathrm{d}t'. \qquad \text{(A.30)}$$

Thus, the description with the stochastic time delays is consistent, and it is sufficient to consider the descendants of the first mutant only. Since early occurrences are extremely rare they will hardly ever be seen in the simulations or in the experiments; however they contribute very strongly to the overall average since early mutations (jackpots) can have an enormous number of descendants.

Even if the description is consistent accounting only for the descendants of the first i-mutant to appear, others will contribute also. In fact, the average number of mutants from $i-1$ to i that contributes descendants to the new variant is given by

$$F_i = 1+\int_{T_i}^{\infty} u_i N_{i-1}(t) q_s^{\infty}(t) \approx 1+\frac{u_i N_T}{s_i+D}\,\ln\!\left(\frac{1}{g_i^0}\right). \qquad \text{(A.31)}$$

This gives the expected number of founders contributing to the new adaptive variant i. If $u_i N_T$ is large, there will be a large number of mutants arising and contributing to the population of the new variant.

Growth of a Neutral Variant

Let us again consider the case where the adaptive variant $i-1$ has become established at time 0 and follow the growth of the new variants from here. The neutral variant $(i-1,\,w)$ that derives from variant $i-1$, will behave similarly to variant i discussed above. However, since the growth advantage relative to $i-1$ is zero rather than s_i, other approximations are required for some of the quantities, notably the survival probability. A neutral variant that occurs early, before the $i-1$ variant has taken over the population, will start effectively with a growth advantage s_{i-1}, and will grow initially like a selected one in an exponential inflation phase; mutants that appear late will simply experience the random drift of a truly neutral variant.

The excess growth rate for the neutral variant is the same as that for its parent as given by eqn (A.17) above. The survival probability at time t_s for a neutral variant derived from $i-1$ at time t' can be calculated from eqn (A.5)

$$q_s^{(w)}(t_s;\,t') = \frac{g_{i-1}^0+\exp(-s_{i-1}t')}{g_{i-1}^0[1+D(t_s-t')]+\exp(-s_{i-1}t')}, \times\left[1+\frac{D}{s_{i-1}}\,(1-\exp(-s_{i-1}(t_s-t')))\right] \qquad \text{(A.32)}$$

if it occurs at some time t' before the presence of variant i has become significant. Thus the average number of descendants at time t conditional on survival until time t can be calculated as

$$\langle n_{i-1,w}(t;\,t')\rangle_s = \frac{\langle n_{i-1,w}(t;\,t')\rangle}{q_s^{(w)}(t;\,t')}. \qquad \text{(A.33)}$$

Using eqns (A.2) and (A.17) to calculate $\langle n_{i-1,w}(t;\,t')\rangle$, one finds the average fraction of the neutral variant present in the $i-1$ population at time t:

$$f_{i-1,w}(t;\,t') = \frac{\langle n_{i-1,w}(t;\,t')\rangle_s}{N_{i-1}(t)} = \frac{g_{i-1}^0 s_{i-1}[1+D(t-t')]}{N_T g_{i-1}^0 s_{i-1}} = \frac{+(D+s_{i-1})\exp(-s_{i-1}t')-D\,\exp(-s_{i-1}t)}{N_T g_{i-1}^0 s_{i-1}}. \qquad \text{(A.34)}$$

This is the presence based on the first surviving mutation to appear. For large t, but not so large that the linear dilution term contributes significantly, this gives:

$$f_{i-1}^0 = \frac{1+D/s_{i-1}}{N_T g_{i-1}^0}\,\exp(-s_{i-1}t'). \qquad \text{(A.35)}$$

This expresses the average amount of growth for the first surviving mutant during the exponential inflation phase before its parent $i-1$ has taken over. Since this growth is highly stochastic with the distribution as given by eqn (A.9), the stochastic factor, S, from eqn (8) of the main text should be multiplied to eqn (A.35). After the first surviving appearance, mutant variants continue to appear. This continued appearance can be assumed to be deterministic [eqn (6b) of the main text] giving

$$f_{i-1,w}(t;\,t') = S\cdot f_{i-1}^0 + u_w(t-t'). \qquad \text{(A.36)}$$

The rate of appearance is $u_w N_{i-1}$ and the probability that the first surviving occurrence of the neutral variant has not taken place before time t can be calculated from eqn (A.15) and (A.32), which gives

$$Q_0^{(w)}(t) \approx \exp\left[-\frac{u_w N_T g_{i-1}^0}{s_{i-1}+D} \left(\exp(s_{i-1}t)-1\right) \right]. \tag{A.37}$$

The approximation is good as long as $u_w N_T \gg D$. From this one can calculate the stochastic appearance times $T_{i-1,w}$ in accordance with eqn (9) of the main text.

If the result [eqn (A.36)] is averaged over the distribution of the appearance times, $-\mathrm{d}Q_0^{(w)}(t')/\mathrm{d}t'$, in the same way as eqn (A.30), one finds

$$f_{i-1}(t) = u_w t, \tag{A.38}$$

as expected from the average of the stochastic equations [eqn (A.13)] or from the deterministic ones. Numerically, one also finds that the variance from eqn (A.14) holds for a large range of reasonable parameter values. This verifies the consistency of the description using eqn (A.36) to describe the stochastic growth of a neutral variant.

If $u_w N_T$ is large, the appearance times are more narrowly distributed around the average (cf. Fig. 1)

$$\bar{T}_{i-1,w} = \frac{1}{s_{i-1}} \ln\left[\frac{D+s_{i-1}}{u_w N_T g_{i-1}^0} \right] \tag{A.39}$$

and the initial condition is approximately

$$f_{i-1}^0 = \frac{u_w}{s_{i-1}}. \tag{A.40}$$

Most of the appearance will be clustered around these averages [eqns (A.39–40)] but occasional jackpots will appear early and contribute an enormous amount of descendants; these rare occurrences with large contributions will make the overall average conform to eqn (A.38).