Current Biology, Vol. 12, 1040–1045, June 25, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S0960-9822(02)00896-5

Fitness Effects of Fixed Beneficial Mutations in Microbial Populations

Daniel E. Rozen,[1,2] J. Arjan G.M. de Visser,[1,3] and Philip J. Gerrish[4,5,6]
[1]Center for Microbial Ecology
Michigan State University
East Lansing, Michigan 48824
[2]MSI/WTB Complex
University of Dundee
Dundee DD1 5EH
United Kingdom
[3]Department of Genetics
Wageningen University
Arboretumlaan 4
6703 BD Wageningen
The Netherlands
[4]Theoretical Biology and Biophysics
Los Alamos National Laboratory
Los Alamos, New Mexico 87545
[5]Programa de Investigación
en Matemáticas Aplicadas
Instituto Mexicano del Petróleo
Eje Central Lázaro Cárdenas Number 152
Colonia San Bartolo Atepehuacán
México, D.F. 07730
México

Summary

Beneficial mutations are intuitively relevant to understanding adaptation [1–3], yet not all beneficial mutations are of consequence to the long-term evolutionary outcome of adaptation. Many beneficial mutations—mostly those of small effect—are lost due either to (1) genetic drift [4, 5] or to (2) competition among clones carrying different beneficial mutations, a phenomenon called the "Hill-Robertson effect" for sexual populations [6] and "clonal interference" for asexual populations [7]. Competition among clones becomes more prevalent with increasing genetic linkage and increasing population size, and it is thus generally characteristic of microbial populations [8, 9]. Together, these two phenomena suggest that only those beneficial mutations of large fitness effect should achieve fixation, despite the fact that most beneficial mutations produced are predicted to have very small fitness effects [10, 11]. Here, we confirm this prediction—both empirically and theoretically—by showing that fitness effects of fixed beneficial mutations follow a distribution whose mode is positive.

Results

Theory

Figure 1 is presented as an intuitive guide for the following theoretical developments. We here distinguish between three kinds of advantageous mutations: (1) beneficial mutations, all mutations that increase fitness, (2) contending mutations, the subset of beneficial mutations that are not lost by drift and are thus viable contenders for fixation, and (3) fixed beneficial mutations, the subset of contending mutations that achieve fixation. In what follows, we derive the corresponding fitness effect distributions for each kind of mutation. The cumulative density functions (cdfs) are denoted $F(s)$, $G(s)$, and $H(s)$, respectively. The derivations of $F(s)$ and $G(s)$ are not new ($F(s)$ is found in [11–13]; $G(s)$ is implicit in [5, 14] and given in [15]), but $H(s)$ is a novel contribution.

Fitness Effects of Beneficial Mutations

All mutant genotypes produced in the population may be considered to have fitnesses drawn from some unknown distribution, $p(w)$ (Figure 1A). [Note that $p(w)$ may be dependent on the currently dominant genotype, meaning that it can be subject to change with each fixation event. See [4] for a description of how the right tail—the relevant part—of $p(w)$ should change with fixations, under the assumption of Fisherian fitness landscapes. These changes in the tail of $p(w)$ translate to changes in a single parameter, α, as defined below.] Suppose we know, in advance, that a large number of genotypes, n, will be sampled by mutation during the time required for the appearance and fixation of a beneficial mutation. At the outset, the wild-type is the variant of highest fitness, but it is not the best of n genotypes, as a large number of mutants have not yet been sampled. Suppose that the wild-type at the outset is the genotype of fitness rank i. That is, the initial wild-type is the i^{th} most fit of the n possible genotypes. Then, the genotypes of fitness rank 1, 2, 3, …, $i - 1$ are all potential beneficial mutants. Orr [16] used extreme value theory to show that the fitness effect of the next beneficial mutation, given that it can result in a genotype of fitness rank 1, 2, 3, …, or $i - 1$, is exponentially distributed. And, surprisingly, the single parameter for this distribution is independent of i, meaning that the fitness effect distribution does not depend on the "starting point" fitness of the wild-type. (Orr assumes that the parent fitness distribution, $p(w)$, has an exponential tail. Based on this assumption, he finds that fitness differences have an exponential distribution that is independent of i. This is different from our claim that selection coefficients have such an invariant exponential distribution. We defend our application of Orr's result, however, on the grounds that in our experiments and for the purposes of our theory, the wild-type fitness is known and is the same for all replicate populations, such that fitness differences and selection coefficients differ only by the scaling factor $k = 1/W_i$, where W_i is wild-type fitness. Furthermore, we have found that under a different but equally encompassing assumption about the tail behavior of $p(w)$, selection coefficients and not fitness differences have an invariant exponential distribution.) Because of the typically large sizes of microbial populations, the number n can be assumed to be large, and because the overwhelming majority of mutations are not beneficial, i can be assumed small relative to n. Thus, the asymptotic require-

[6]Correspondence: pgerrish@yahoo.com

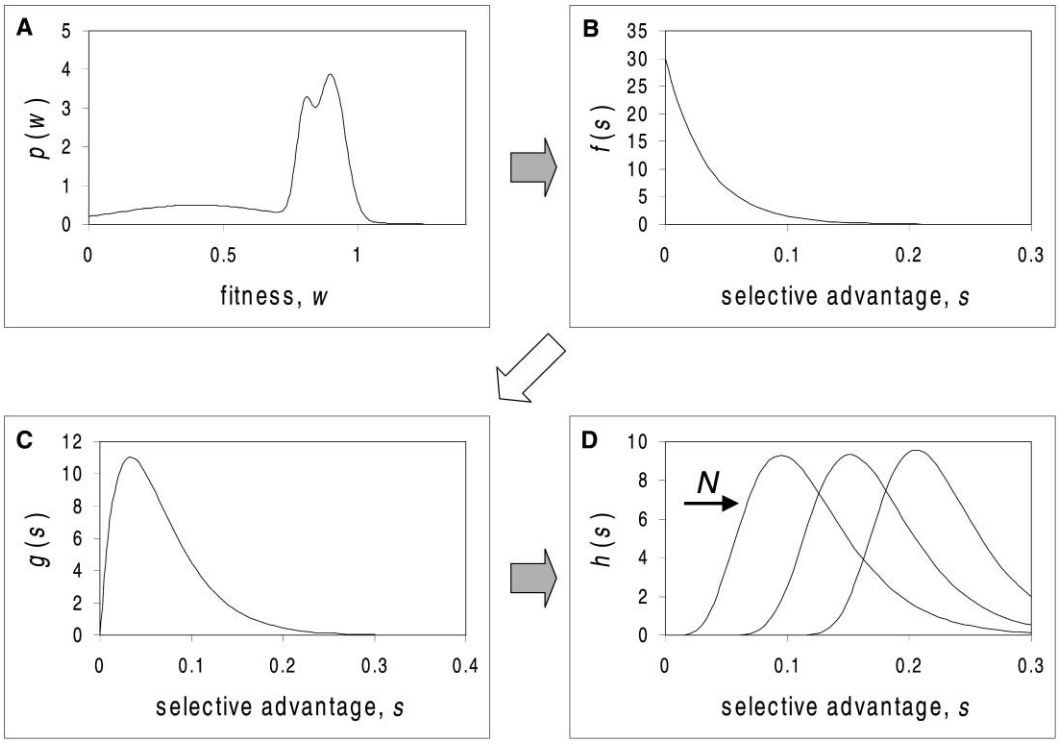

Figure 1. Theoretical Predictions

(A) A plausible but arbitrary fitness distribution of mutations produced in the population, $p(w)$.

(B) Independent of $p(w)$, the fitness effects of beneficial mutations are exponentially distributed [11–13] with density $f(s)$.

(C) The fitness-effect distribution for contending mutations, $g(s)$. This distribution is skewed to the right of $f(s)$ because beneficial mutations of large effect are more likely to survive genetic drift than beneficial mutations of small effect.

(D) Distributions for fitness effects of fixed beneficial mutations, $h(s)$, at three different population sizes. Because fixed beneficial mutations are the best of several contending mutations, the $h(s)$ are distributions of maxima of samples drawn from $g(s)$ and are thus shifted to the right of $g(s)$. Therefore, $h(s)$ propagates to the right as population size increases, as indicated by the arrow. Note that the distribution hardly changes shape as it propagates, indicating that it quickly converges to its asymptotic form. The gray arrows indicate very robust transforms (meaning that the general form of the "output" is highly independent of the "input"), whereas the clear arrow indicates a less robust transform.

ments are met [16], and the fitness effects of beneficial mutations can be assumed to have the exponential density $f(s) = \alpha e^{-\alpha s}$ (Figure 1B), with corresponding cdf, $F(s) = 1 - e^{-\alpha s}$, where $1/\alpha$ is the average selective advantage of beneficial mutations. This quantitative prediction agrees qualitatively with Fisher's argument [10] that beneficial mutations of smaller effect should always be more abundant than those of larger effect.

Fitness Effects of Contending Mutations

A significant fraction of all beneficial mutations produced will become extinct simply due to random fluctuations in frequency, or "genetic drift." These mutations typically do not attain high frequency and are thus of little consequence from a long-term evolutionary standpoint. The beneficial mutations that are *not* lost by drift are here called "contending mutations," and, on average, their selective advantages are larger than those of beneficial mutations. For any organism in almost any environmental regime, the probability of surviving drift can be shown to be approximately Ks, where K is a constant [5, 17, 18]. Therefore, the probability density of fitness effects for beneficial mutations that survive drift is $g(s) = sf(s)/\int_0^\infty uf(u)du = \alpha^2 s e^{-\alpha s}$ (Figure 1C), a gamma density [15, 19] with shape parameter equal to 2 and with corresponding cdf, $G(s) = 1 - (1 + \alpha s)e^{-\alpha s}$.

Fitness Effects of Fixed Beneficial Mutations

Of several contending mutations produced, only one will achieve fixation. This will be the most fit of the several contending mutations produced. If it is known a priori that each fixed beneficial mutation will be the best of, say, n contending mutations, then the cdf for fitness effects of fixed beneficial mutations is simply $[G(s)]^n$ (e.g., see [20]). Because we do not know n a priori, we compute this cdf as

$$\sum_{n=0}^{\infty} p_n[G(s)]^n,$$

where p_n is the probability that the fixed beneficial mutation will be the best of n contending mutations. Put differently, fitness effects of fixed beneficial mutations have cdf $\tilde{P}(G(s))$, where $\tilde{P}$ is the probability generating function (pgf) for numbers of contending mutations from which a fixed beneficial mutation arises. This pgf has been derived by Gerrish [21] and depends on a parameter, j, the number of contending mutations that arise between the appearance of the beneficial mutation destined for fixation and its fixation. This number will be some function of the fitness effect of the beneficial mutation in question: as this fitness effect decreases, for

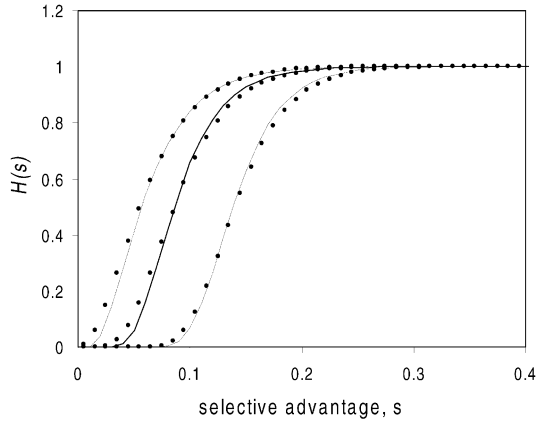

Figure 2. Comparing Approximate $H(s)$ from Equation 1 (lines) with Simulation Results (dots)

Parameter values are $N = 3 \times 10^7$, $\alpha = 35$, and the three curves represent three different beneficial mutation rates, $\mu = 2 \times 10^{-10}$, $\mu = 2 \times 10^{-9}$, and $\mu = 2 \times 10^{-8}$. Simulations are much simpler than those shown in Figure 3 and described in the Experimental Procedures. Here, fitness effects for beneficial mutations, S, are drawn at random from an exponential distribution with parameter α, and time intervals are drawn at random from an exponential distribution with parameter μN. Beneficial mutations survive drift, i.e., become contending mutations, with probability $4S$. Each new "best" contending mutation (with selective advantage S_{MAX}) resets the clock. If no superior contending mutation appears in the subsequent $\ln(N/2)/S_{MAX}$ generations, then the mutation with selective advantage S_{MAX} is considered fixed, and the process is stopped. Fifty thousand simulations were run for each mutation rate.

example, the time between appearance and fixation will increase, thus allowing time for an increasing number of subsequent contending mutations to appear. We thus replace the parameter j with a function, $\lambda(s)$, giving the approximate cdf for fitness effects of fixed beneficial mutations:

$$H(s) \approx \tilde{P}(G(s)|\lambda(s)) \quad (1)$$

where $\tilde{P}$ is given in [21], and corresponding density, $h(s) = H'(s)$ (Figure 1D). We compare this approximation with simulation results in Figure 2. The practical implementation of this equation is complicated by the fact that $\lambda(s)$ is a continuous function, whereas j is a discrete parameter. As a heuristic remedy for this dilemma, it can be shown that the pgf given in [21] has an equivalent expression that here takes the form

$$\tilde{P}(G(s)|\lambda(s)) = 1 - \exp\{-\int_0^{G(s)} \frac{u^{\lambda(s)}}{1-u}\, du\} \quad (2)$$

which is continuous in $\lambda(s)$. Alternatively, an asymptotic expression for Equation 1 may be employed for the case of large population size, N, and/or high beneficial mutation rate, μ. Such an expression is derived in the Experimental Procedures; it is

$$H(s) \xrightarrow{\mu N} e^{-e^{\gamma}\lambda(s)[1-G(s)]}. \quad (3)$$

where $\gamma \approx 0.577$, Euler's constant. It should be noted that Equation 3 is an "extreme value distribution" [22, 23], which is to be expected given that clonal interference insures the fixation of only the best mutants. To evaluate Equations 1 or 3, it is necessary to have expressions for $G(s)$ and $\lambda(s)$, which are derived above and in the Experimental Procedures, respectively. It is revealed in the Experimental Procedures that, for the purposes of our experiment, the only unknown parameters are beneficial mutation rate, μ, and the exponential parameter for the fitness effect distribution of beneficial mutations, α.

To summarize, each transformation accounted for above has the effect of (1) increasing average fitness effect and (2) increasing the independence of the distribution on the parent fitness distribution $p(w)$. The final result is a bell-shaped distribution, $h(s)$, for fitness effects of fixed beneficial mutations, whose mean is perhaps surprisingly large and whose general shape does not depend on the unknown shape of the parent fitness distribution.

Experiment

To generate a library of *E. coli* genotypes that each contained a single fixed beneficial mutation, we employed a technique that enabled us to collect fixed beneficial mutations as they arose. Initially, a specified fraction of each experimental population carried a neutral marker. After a period of in vitro evolution, a sudden deviation in marker frequency would indicate that either the marked or unmarked genotype was linked to a beneficial mutation on its way to fixation, i.e., it would indicate a periodic selection event [24]. Following experimental evolution, the fitness of each genotype was determined relative to the unevolved ancestral genotype by placing evolved and ancestral genotypes in short-term direct competition [25].

Experimental Evolution

Two genotypes, isogenic except for a single neutral genetic marker that caused colonies to appear either red (r) or white (w) on indicator agar [26], were placed in mixed culture in test tubes with a minimal glucose medium and serially transferred daily. Thirty replicate evolving populations were monitored for up to 400 generations (60 days) by periodically plating samples on indicator agar to determine the frequency of each marker. Five initial marker ratios ranging from $w/r = 0.01$ to $w/r = 100$ were examined, and total effective population size was $N = r + w \approx 3 \times 10^7$. The spread of an individual beneficial mutation destined for fixation could be observed because it would cause a deviation from the initial ratio, w/r, as one of the genetic backgrounds (either r or w) hitchhiked with the beneficial mutation to high frequency [27, 28]. The subpopulation that was found to increase in frequency was deemed the "winner."

Beneficial mutations were the most likely cause of the observed deviations, as the probability that the deviations were due to drift alone was much less than 0.001 for $w/r = 1$ and less than 0.05 for $w/r = 0.01$ or $w/r = 100$ (based on Crow and Kimura [29]). It is possible, however, that a few of the observed deviations were caused by transiently common beneficial mutations that would in fact ultimately be unsuccessful, i.e., displaced by a superior competitor, if time were allowed (a phenomenon that has been dubbed the "leapfrog" [7]). We

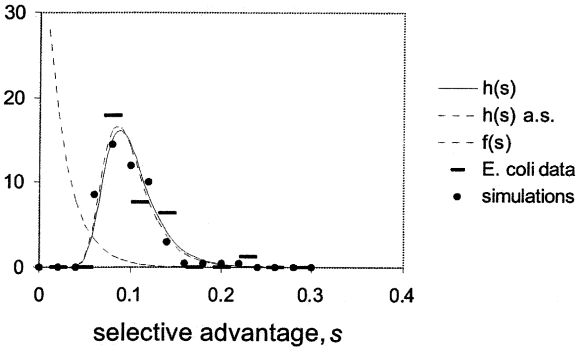

Figure 3. Distributions of Fitness Effects

The dashes represent a histogram generated from fitness measurements of beneficial mutants in the *E. coli* populations. The solid line shows the probability density, $h(s)$ from Equation 1, given maximum likelihood parameter estimates (see Figure 4). The dashed line close to the solid line is the asymptotic expression for $h(s) = H'(s)$, where $H(s)$ is given by Equation 3. The other dashed line (monotonically decreasing) shows the projected underlying distribution for beneficial mutations, $f(s)$. The dots represent a histogram generated from fitness effects of fixed beneficial mutations in 100 simulated bacterial populations (see Experimental Procedures).

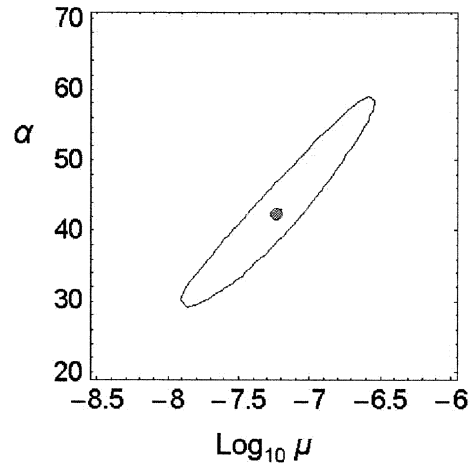

Figure 4. The Oval-Shaped Curve Defines the 95% Confidence Region

This region is computed as the contour line satisfying

$$L(\mu,\alpha) = L(\hat{\mu},\hat{\alpha})\exp\{-\chi^2_{95,1}/2\},$$

where $L(\mu,\alpha)$ is the likelihood function defined in Experimental Procedures, $\chi^2_{95,1}$ is the 95th percentile of the chi-square distribution with one degree of freedom (because only one quantile is being reported), and a hat indicates maximum likelihood estimate. The dot in the middle of the plot is where the maximum likelihood occurs, corresponding to maximum likelihood parameter estimates of $\hat{\mu} = 5.9 \times 10^{-8}$ and $\hat{\alpha} = 42.5$.

do not consider this effect to present a serious difficulty, however, because (1) it might be expected to have occurred in only a few of the replicate populations (see [7] and simulations in the Experimental Procedures), and (2) such transiently common mutations have fitness effects similar to beneficial mutations that achieve fixation, suggesting that the fitness-effect distributions should be little affected by this phenomenon. Another potential confounding factor was the possibility that, in some of the populations, winners had fixed more than one beneficial mutation. Simulation results, however, suggest that this potential complication was in fact very unlikely to occur (see Experimental Procedures).

Fitness Assays

Random colonies from winning subpopulations were picked following experimental evolution. Fitness was estimated as outlined by Lenski et al. [25]. Briefly, clonal samples of each evolved genotype and its ancestor were removed from a −80°C freezer and grown to saturation in the same experimental environment used to isolate mutants. Equal densities of both competitors were then mixed, and the change in their relative densities was measured over the course of 2 days. Relative fitness was calculated as the ratio of the estimated growth rate of each competitor to its ancestor. Fitness of each mutant was determined in four complete blocks.

Fitnesses measured are plotted as a histogram in Figure 3. Here, it is apparent that the fitness effects of fixed beneficial mutations have a bell-shaped distribution. Put differently, fixed beneficial mutations are more likely to have intermediate fitness effects than either small or large fitness effects. This finding is qualitatively in accordance with theoretical predictions.

Discussion

Beneficial mutations are very rare events and are thus difficult to observe. Microbial populations seem ideal for studying beneficial mutations because microbes have short generation times and their populations are large. Microbial populations propagated in the laboratory over a relatively short period of time can undergo billions of replications. In such experiments, beneficial mutations are sure to arise. Beneficial mutations only become detectable, however, when they have achieved observable frequencies in the population. To achieve observable frequency in a population, a beneficial mutation must survive both drift and, to a large degree, clonal interference. Together, both of these "obstacles" select for beneficial mutations of large effect, thus creating a bias in the beneficial mutations that are observed. In our experiments, most beneficial mutations achieved a frequency of 0.5 or greater before they were observed, and therefore, most of the beneficial mutations we observed were probably destined for evolutionary success (see [7] and simulations described in the Experimental Procedures). Thus, what we have characterized here are not beneficial mutations but fixed beneficial mutations. As predicted theoretically, our experimentally derived fitness distribution for fixed beneficial mutations, $h(s)$, tends toward larger values.

Given our equation for $H(s)$ as a function of $G(s)$, it is tempting to suppose that fitness effects of contending mutations, or even beneficial mutations, can be directly inferred (i.e., without making a priori assumptions about the shapes of these distributions) from our fitness effect data for fixed beneficial mutations by computing the inverse of $\tilde{P}(G(s)|\lambda(s))$. Practically, however, this is a very difficult if not intractable problem. Aside from being a difficult problem mathematically, almost no statistical power can be expected. The reason is that many differ-

ent $G(s)$ can give rise to statistically identical $H(s)$. [From a statistical standpoint, it could be said that the mapping from $G(s)$ to $H(s)$ is not one to one.] This is because $H(s)$ is a "distribution of maxima" of "competing" values drawn from $G(s)$. It is well known [23] that a distribution of maxima is only very weakly dependent on the general shape of the parent distribution from which competing values are drawn.

To make inference with respect to the underlying distributions, therefore, we had to appeal to theoretical predictions. These predictions drastically reduce the degrees of freedom involved, thus permitting reasonable inference of the two key parameters. Our study of advantageous mutations thus depends equally on experimental and theoretical results, and it is with this in mind that we decided to present them together.

The inability to directly infer $G(s)$ from $H(s)$, while somewhat frustrating, has the intriguing implication that $H(s)$ is highly independent of the parent fitness distribution. In fact, $H(s)$ results from *two* very robust transforms, as indicated in Figure 1. Furthermore, the theoretical prediction for this distribution is supported by agreement with both empirical and simulated data, as displayed in Figure 3.

Our findings contrast with and complement the results of another recent study of advantageous mutations in *E. coli* populations [3]. While their study did not distinguish between the different kinds of advantageous mutation, according to our definitions they measured fitness effects of contending mutations. These were measured relative to the population fitness at the time of measurement by back-calculating fitness from the change in marker frequency. Measured in this way, fitness effects of contending mutations were found to have a monotonically decreasing (exponential-like) distribution. A factor that may have affected the general shape of their distribution is that population fitness at the time of measurement (their reference point) may have been affected by other contending mutations in the population that were also observed to be at high frequency. Their distribution should thus be skewed to the left of its true shape due to this bias. Without this bias, we contend that their data should follow the distribution of contending mutations, as in our Figure 1C.

Consistent with theoretical predictions, our empirical data in Figure 3 show that fitness effects of fixed beneficial mutations do not have a monotonically decreasing distribution, as is predicted for beneficial mutations, but instead their distribution is centered on mutations with larger fitness effects. This observation is particularly relevant for our understanding of adaptive evolution, because it is the realized outcome of selection that drives the process of adaptation. Specifically, it suggests a more prominent role for large-effect mutations in adaptation, despite the apparent rarity of such mutations. The data and theory presented here characterize the fitness effects of fixed beneficial mutations. In a previous study [21], the timing of fixations of beneficial mutations was characterized. Together, these two studies lay some groundwork for a general model of adaptive evolution in microbial populations.

Experimental Procedures

Number of Subsequent Contending Mutations, $\lambda(s)$

Beneficial mutations may be assumed to grow in frequency logistically, for example, such that the decline in the number of wild-type, x, is given by the equation $dx/dt = -sx(1 - x/N)$, where N is population size. (There are different models of growth that one could choose to describe the growth of the beneficial mutant lineage and the corresponding decline of the wild-type. We have chosen a common model, that of logistic growth, but it is only one model. In addition to having several growth models to choose from, many levels of detail can be added to refine one's model.) (If the population size fluctuates, reasonable approximations may be obtained by defining N to be the fixation effective population size given in [30], and replacing K below with $KN/\overline{N}$, where $\overline{N}$ is the arithmetic mean population size [17].) The total number of replications of wild-type between the appearance of the beneficial mutation destined for fixation and its fixation is thus

$$\int_0^{t_f} x(t)dt = \int_{N-1}^{(1-\phi)N} \frac{dx}{-s(1 - x/N)} = \frac{N \ln(\phi N)}{s},$$

where t_f denotes the number of generations required for fixation (not needed), and ϕ is the frequency at which a mutation is deemed fixed. Each of these replication events has probability μ of producing a beneficial mutation, and on average each beneficial mutation produced has approximate probability

$$K \int_0^{\infty} uf(u)du = K/\alpha$$

of not being lost by drift. The number of subsequent contending mutations is thus $\lambda(s) = K\mu N \ln(\phi N)/(\alpha s)$. The parameter K is shown in [31] to be approximately equal to 2 for bacteria in which mutations occur during DNA replication and 2.8 when mutations occur at random times during replication or growth. The fixation-effective and average population sizes are known for our experiments to be $N \approx 3.3 \times 10^7$ and $\overline{N} \approx 1.1 \times 10^8$, respectively, and making the reasonable assumption that mutations occur during replication, we thus take $KN/\overline{N} = 0.6$. For our purposes, therefore, the only unknown parameters in Equation 1 are μ and α.

Asymptotic Distribution for Fitness Effects of Fixed Beneficial Mutations

Clonal interference insures fixation of only the best of several contending mutations. Thus, it should be only the tail probabilities of $G(s)$ that affect $H(s)$. This a priori observation suggests that, for the purposes of approximation, the relevant region to consider is that of relatively large s and $G(s)$ close to 1. In this region, Taylor series expansion of $\tilde{P}$ as given by [21] provides

$$\ln[H(s)] \approx [G(s) - 1] \exp\{\sum_{k=1}^{\lambda(s)} \frac{[G(s)]^k}{k}\}. \quad (4)$$

Constraint to the region of large N or large μ [implying large $\lambda(s)$] further provides

$$\sum_{k=1}^{\lambda(s)} \frac{[G(s)]^k}{k} \rightarrow \ln[\lambda(s)] + \gamma,$$

where $\gamma \approx 0.577$, Euler's constant. Thus,

$$\ln[H(s)] \xrightarrow{\mu N} -e^{\gamma} \lambda(s)[1 - G(s)] \quad (5)$$

This gives an asymptotic expression for the distribution function for fitness effects of fixed beneficial mutations, given in the main text as Equation 3.

Simulations

We simulated our experiments to evaluate whether our criteria for calling a mutation successful corresponded to true fixed beneficial mutations. The simulations started with homogeneous fitness and two equally represented subpopulations (red and white), and they tracked each lineage created by beneficial mutation. A lineage started with a single mutant individual and grew stochastically until

it acquired 100 members, after which its growth was deterministic. (Many lineages quickly died out due to their stochastic trajectory, or genetic drift, and thus never acquired 100 members.) Let f_i denote the frequency of lineage i, and let μ denote beneficial mutation rate. Each generation, a Poisson distributed number of new beneficial mutations arose in each lineage with mean μf_i, forming new lineages. The new beneficial mutations conferred selective advantages that were drawn from an exponential distribution. The parameter for this distribution, α, as well as the beneficial mutation rate, μ, were taken to be equal to the analytical estimates presented in the caption of Figure 3. When a significant deviation in frequencies occurred, the subpopulation that had the higher frequency was deemed the winner, and the fitnesses of both subpopulations were recorded. To check for multiple fixations, we tracked the number of fixed beneficial mutations accumulated by the winner. We found this number to be greater than 1 in 0 out of 100 trials. To check our criteria for identifying fixed beneficial mutations, we allowed the simulations to continue to run for 2000 generations after a putative fixation was identified and checked that the lineage of highest frequency at this later time carried the putative fixed mutation. In 21 out of 100 trials it did not, suggesting that roughly 6 of the 30 putative fixations that we identified in the *E. coli* populations would in fact be ultimately unsuccessful.

Maximum Likelihood Estimation of Parameters

Let s_i denote the selective advantage conferred by the winner in the i^{th} replicate population. Write $h(s_i,\phi_i|\mu,\alpha)$ to denote the value of the probability density, $h(s) = H'(s)$, evaluated at measured values s_i, and assumed parameter values μ and α. The value N_i is the initial size of the winning subpopulation in the i^{th} replicate population, and we defined $\phi_i = N_i/N$, because a significant deviation in the initial frequency was when a beneficial mutation was deemed fixed. The maximum likelihood estimates of beneficial mutation rate, μ, and exponential parameter, α, were determined by finding the values for these two parameters that maximized the function

$$L(\mu,\alpha) = \prod_{i=1}^{28} h(s_i,\phi_i|\mu,\alpha).$$

(Estimates are given in the caption of Figure 4.) We indicate only 28 data points in this equation, because the selective advantages measured in two of the replicate populations were not large enough to allow time for their fixation, and marker deviations in these populations were suspiciously small. We note that theory and simulations strongly suggest that the most probable explanation for no detectable fixation in these two populations is *not* that its fitness effect was too small but simply that there was not enough time allowed for a fixation to occur. Yet, even in the unlikely event that mutations of very small effect were somehow fixed in these two lines, the inclusion of these two extra data points would still not change the bell-shape of the histogram shown in Figure 3.

Acknowledgments

We thank R. Lenski, A. Orr, two anonymous reviewers, M. Imhof, T. Cooper, F. Moore, V. Cooper, N. Hajela, and L. Ekunwe for comments, fruitful discussions, and help with lab work. Support was received from a National Science Foundation grant to R. Lenski, a fellowship from the Netherlands Organization for Scientific research to J.A.G.M.d.V., and a fellowship from Los Alamos National Laboratory to P.J.G.

Received: December 20, 2001
Revised: April 26, 2002
Accepted: April 29, 2002
Published: June 25, 2002

References

1. Travisano, M. (2001). Towards a genetical theory of adaptation. Curr. Biol. *11*, R440–R442.
2. Bataillon, T. (2000). Estimation of spontaneous genome-wide mutation rate parameters: whither beneficial mutations? Heredity *84*, 497–501.
3. Imhof, M., and Schlotterer, C. (2001). Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. Proc. Natl. Acad. Sci. USA *98*, 1113–1117.
4. Orr, H.A. (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution *52*, 935–949.
5. Haldane, J.B.S. (1927). The mathematical theory of natural and artificial selection, part V: Selection and mutation. Proc. Camb. Philol. Soc. *23*, 838–844.
6. Hill, W.G., and Robertson, A. (1966). The effect of linkage on limits to artificial selection. Genet. Res. *8*, 269–294.
7. Gerrish, P.J., and Lenski, R.E. (1998). The fate of competing beneficial mutations in an asexual population. Genetica *102–103*, 127–144.
8. de Visser, J.A.G.M., Zeyl, C.W., Gerrish, P.J., Blanchard, J.L., and Lenski, R.E. (1999). Diminishing returns from mutation supply rate in asexual populations. Science *283*, 404–406.
9. Miralles, R., Gerrish, P.J., Moya, A., and Elena, S.F. (1999). Clonal interference and the evolution of RNA viruses. Science *285*, 1745–1747.
10. Fisher, R.A. (1930). The Genetical Theory of Natural Selection (Oxford: Oxford University Press).
11. Gillespie, J.H. (1991). The Causes of Molecular Evolution (Oxford: Oxford University Press).
12. Smid, B., and Stam, A.J. (1975). Convergence in distribution of quotients of order statistics. Stochastic Processes and their Applications *3*, 287–292.
13. Shorrock, R.W. (1972). On record values and record times. J. Appl. Prob. *9*, 316–326.
14. Kimura, M. (1983). The neutral theory of molecular evolution (Cambridge: Cambridge University Press).
15. Otto, S.P., and Jones, C.D. (2000). Detecting the undetected: Estimating the total number of loci underlying a quantitative trait. Genetics *156*, 2093–2107.
16. Orr, H.A. (2002). An invariance rule for fitness effects of beneficial mutations. Genetics, in press.
17. Ewens, W.J. (1967). The probability of survival of a new mutant in a fluctuating environment. Heredity *22*, 438–443.
18. Ewens, W.J. (1969). Population Genetics (London: Methuen Press).
19. Wahl, L.M., Gerrish, P.J., and Saika-Voivod, I. (2002). Evaluating the impact of population bottlenecks in experimental evolution. Genetics, in press.
20. Feller, W. (1968). An Introduction to Probability Theory and Its Applications, Third edition (New York: John Wiley & Sons, Inc.).
21. Gerrish, P.J. (2001). The rhythm of microbial adaptation. Nature *413*, 299–302.
22. Fisher, R.A., and Tippett, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Camb. Philol. Soc. *24*, 180–184.
23. Gumbel, E.J. (1958). Statistics of Extremes (New York: Columbia University Press).
24. Atwood, K.C., Schneider, L.K., and Ryan, F.J. (1951). Periodic selection in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *37*, 146–155.
25. Lenski, R.E., Rose, M.R., Simpson, S.C., and Tadler, S.C. (1991). Long-term experimental evolution in Escherichia coli: I. Adaptation and divergence during 2000 generations. Am. Nat. *138*, 1315–1341.
26. Levin, B.R., Stewart, F.M., and Chao, L. (1977). Resource-limited growth, competition and predation—a model and experimental studies with bacteria and bacteriophage. Am. Nat. *111*, 3–24.
27. Lenski, R.E. (1988). Experimental studies of pleiotropy and epistasis in *Escherichia coli*. Evolution *42*, 433–440.
28. Chao, L., and Cox, E.C. (1983). Competition between high and low mutating strains of *Escherichia coli*. Evolution *37*, 125–134.
29. Crow, J.F., and Kimura, M. (1970). An Introduction to Population Genetics Theory (New York: Harper & Row).
30. Otto, S.P., and Whitlock, M.C. (1997). The probability of fixation in populations of changing size. Genetics *146*, 723–733.
31. Johnson, T., and Gerrish, P.J. (2002). The fixation probability of a beneficial allele in a population dividing by binary fission. Genetica, in press.