

The Distribution of Fitness Effects Among Beneficial Mutations

H. Allen Orr¹

Department of Biology, University of Rochester, Rochester, New York 14627

Manuscript received October 10, 2002

Accepted for publication January 5, 2003

ABSTRACT

We know little about the distribution of fitness effects among new beneficial mutations, a problem that partly reflects the rarity of these changes. Surprisingly, though, population genetic theory allows us to predict what this distribution should look like under fairly general assumptions. Using extreme value theory, I derive this distribution and show that it has two unexpected properties. First, the distribution of beneficial fitness effects at a gene is exponential. Second, the distribution of beneficial effects at a gene has the same mean regardless of the fitness of the present wild-type allele. Adaptation from new mutations is thus characterized by a kind of invariance: natural selection chooses from the *same* spectrum of beneficial effects at a locus independent of the fitness rank of the present wild type. I show that these findings are reasonably robust to deviations from several assumptions. I further show that one can back calculate the mean size of new beneficial mutations from the observed mean size of *fixed* beneficial mutations.

ADAPTATION is a two-step process: (i) alleles having different effects on fitness arise by mutation and (ii) those alleles that improve fitness tend to increase in frequency by natural selection. A good part of classical population genetics focuses on the second step of this process, including calculation of the probability that natural selection will fix a new favorable mutation (HALDANE 1927) and of the rate at which such a mutation will increase in frequency (HALDANE 1924). But the first step in adaptation, the origination of new beneficial mutations, has been less well studied. We cannot say, for instance, if beneficial mutations of small effect are more common than those of large effect and, if so, how much more common.

This is unfortunate as a number of aspects of adaptive evolution depend on the distribution of beneficial fitness effects. For example, the mean increase in fitness that occurs during substitution of a beneficial mutation must depend on the spectrum of effects among new mutations presented to natural selection. Large jumps in fitness, for instance, are possible only if mutations of large favorable effect occur.

The most direct approach to finding the distribution of fitness effects among new beneficial mutations is empirical. But while possible in principle this approach has proved difficult in practice. There are two main problems: beneficial mutations are rare and beneficial mutations of small effect are difficult to detect. Because of this, study of experimental microbial populations would seem to provide the best hope of characterizing the distribution of beneficial effects: the combination of large population size and short generation time means that many beneficial mutations can be sampled

in a short time (LENSKI and TRAVISANO 1994; WICHMAN *et al.* 1999; HOLDER and BULL 2001). But even in microbes it has proved difficult to infer the distribution of beneficial effects. The main reason is that the beneficial mutations seen in most experiments are not a random sample of new mutations but rather those that have escaped stochastic loss. [Most beneficial mutations are accidentally lost when rare; the probability of loss depends on the size of the mutation's fitness effect (HALDANE 1927).] IMHOF and SCHLOTTERER (2001), for instance, recently attempted to characterize the distribution of fitness effects among new mutations in *Escherichia coli*. But because their experimental design [a variation on periodic selection (ARWOOD *et al.* 1951)] depended on detection of favorable alleles that had reached appreciable frequencies, the distribution of effects observed was actually that for those "lucky" mutations that had escaped stochastic loss, not that for new mutations. Similarly, ROZEN *et al.* (2002) recently characterized the distribution of fitness effects among *fixed* beneficial mutations in *E. coli*. This distribution is also not the same as among new beneficial mutations, as ROZEN *et al.* (2002) emphasize. Indeed the distribution of fitness effects among fixed mutations in asexual microbes is distorted by both stochastic loss *and* clonal interference, the competitive exclusion of a mutation of small beneficial effect by one of larger beneficial effect in nonrecombining genomes or chromosome regions (GERRISH and LENSKI 1998; ROZEN *et al.* 2002). Although some experiments have attempted to assay beneficial mutations before they are subject to stochastic loss (BULL *et al.* 2000), these experiments are compromised by another of the problems noted above: they cannot detect beneficial mutations of small effect.

Given these difficulties, it seems worth asking if population genetic theory can provide any insight into the

¹Author e-mail: aorr@mail.rochester.edu

expected distribution of beneficial effects among new mutations. GILLESPIE (1983, 1984, 1991) suggested that the answer is yes. Using extreme value theory, he showed that the fitness gap between high fitness alleles is exponential. If, for instance, we consider a wild-type allele that can mutate to a single beneficial allele, the difference in fitness between these two alleles should be exponentially distributed. This result has been widely cited in the literature (GERRISH and LENSKI 1998; OTTO and JONES 2000; WAHL and KRAKAUER 2000; ORR 2002; ROZEN *et al.* 2002). It has not, however, always been appreciated that Gillespie's result concerns a special case. As OTTO and JONES (2000) emphasize, Gillespie considers only the distribution of fitness differences between "adjacent" alleles, *e.g.*, the difference in fitness between the second-best allele (the present wild type) and the best allele (the beneficial mutant). Gillespie's work thus provides the distribution of fitness effects among new beneficial mutations only if the present wild type can mutate to a *single* beneficial allele. We would obviously like to know the expected distribution of beneficial effects in the general case where the present wild type might mutate to two, or three, or four, etc., different beneficial alleles.

Here I derive this distribution. I show that it has two surprising properties: (i) the distribution of fitness effects among new beneficial mutations is always exponential and (ii) the distribution is invariant; *i.e.*, it has the same mean regardless of the starting fitness rank of the wild-type allele. Our key assumption is that the starting wild-type allele has relatively high fitness.

THE MODEL AND RESULTS

The biological scenario: Following GILLESPIE's (1983, 1984, 1991) "mutational landscape" model, I consider a population that was, until recently, well adapted to the environment. In particular, I consider a population that is essentially fixed for a wild-type sequence that was—until the recent environmental change—the fittest available at the gene. Following the environmental change, the wild type has dropped in fitness. The wild-type sequence can mutate to many alternative sequences. Of these, natural selection is essentially constrained to surveying those that differ from wild type by single-point mutations, as first pointed out by MAYNARD SMITH (1970; see also GILLESPIE 1984). (Double mutations are too rare to be of much evolutionary significance; for the same reason, epistasis among new mutations can be ignored. Our results will hold for small genomes, as well as for single genes, as long as genomes are small enough that most mutations arise singly.) Given a gene that is L bp long, we thus need to consider only the $m = 3L$ single-mutational step mutant sequences. For now, we assume that each of these m mutations arises with equal frequency, reflecting a constant and low per-nucleotide mutation rate. This assumption is relaxed later.

Although we know little about the fitnesses of mutant sequences at any gene, it is clear that most of the m mutations will be less fit than the present wild type. This follows from two facts. First, environments are autocorrelated through time, making it unlikely that the best sequence today will be the worst tomorrow (GILLESPIE 1983). Second, a considerable fraction of mutations are unconditionally lethal or strongly deleterious, making it unlikely that the wild type would fall into the company of such alleles. Given our nearly complete ignorance of mutant fitnesses at most genes, KIMURA (1983) and GILLESPIE (1983, 1984, 1991) suggested that one simply assume that the fitnesses of alternative alleles at a gene are drawn from some probability distribution. Importantly, in this article we do not need to specify this distribution. We assume only that, of the relevant $m + 1$ alleles (m mutations plus wild type), the wild type has relatively high fitness. The wild type can mutate, in other words, to a small number of beneficial sequences.

Distribution of beneficial effects: To find the distribution of fitness effects among those few mutations that are beneficial, we first rank the absolute fitnesses of the m mutant and one wild-type sequences: the fittest allele is given rank 1, the next fittest rank 2, and so on (Figure 1). The wild-type allele has rank i , where i is small. If a typical gene is $L = 1000$ bp long, then $m = 3000$ and i might range from, say, 2 to 25. The fitness gaps between adjacent alleles are labeled Δ_1 , Δ_2 , etc., as shown in Figure 1. Thus a mutation from wild-type allele i to favorable allele $i - 1$ improves absolute fitness by $\Delta W = \Delta_{i-1}$, while a mutation from wild-type allele i to favorable allele 1 improves fitness by $\Delta W = \Delta_{i-1} + \dots + \Delta_2 + \Delta_1$. The overall distribution, $f(\Delta W|i)$, of fitness effects among beneficial mutations when starting from wild-type allele i is the mixed distribution formed by considering all such possibilities. In symbols,

$$f(\Delta W|i) = \frac{1}{i-1} \sum_{j=1}^{i-1} f(\Delta W|i, j), \quad (1)$$

where $f(\Delta W|i, j)$ is the probability density of fitness effects when mutating from an allele of rank i to a beneficial allele of rank j .

Equation 1 shows that if we knew $\Delta_1, \Delta_2, \dots, \Delta_{i-1}$ we would know the distribution of fitness effects among beneficial mutations. Although we have not specified the distribution of allelic fitnesses, we can, surprisingly, still say something about these fitness spacings. The reason, as GILLESPIE (1983) first saw, is that adaptation is confined to the right-hand (fittest) tail of the distribution of allelic fitnesses. This fact lets us take advantage of certain limiting results from extreme value theory that describe the behavior of the top several draws from any reasonable distribution. Formally, the distributions we consider belong to the so-called Gumbel type, a broad category that includes most "ordinary" distributions like the normal, lognormal, exponential, gamma,

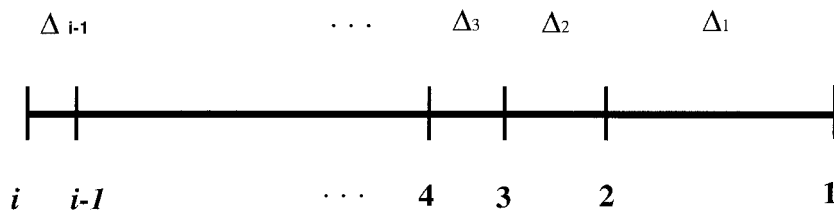


FIGURE 1.—The fitness ranks of the top several alleles. In total $m + 1$ alleles (m mutant alleles plus the wild type) are drawn from an unknown distribution of allelic fitnesses and ranked. The fittest allele has rank 1 and the wild-type allele has rank i . The spacings between adjacent alleles, Δ , represent differences in absolute fitness. These spacings grow smaller as one moves toward the median allele, as shown.

Weibull, logistic, etc. Roughly speaking, this class excludes only exotic distributions like the Cauchy (which has no moments) and many (though not all) distributions that are bounded on the right (GUMBEL 1958; GILLESPIE 1983, 1984, 1991). The APPENDIX provides mathematical details. Although most extreme value theory holds asymptotically as the number of draws from a distribution approaches infinity, the fact that a wild-type sequence can mutate to a very large number of alternate sequences suggests that these asymptotic results should provide good approximations.

For our purposes, the most important of these limit theorems describes the spacings, Δ_j , between the top several draws, *i.e.*, the fittest several alleles. Although for any particular wild-type sequence and set of mutants the Δ_j 's are constants, these extreme spacings will in general be random variables. Extreme value theory shows that these Δ_j 's are asymptotically independent exponentially distributed random variables *regardless* of the distribution of allelic fitnesses. Theory also shows that these spacings grow smaller as one moves toward the median allele as shown in Figure 1. In particular, $E[\Delta_j] = E[\Delta_1]/j$, where the constant $E[\Delta_1]$ depends on the form of the distribution of allelic fitnesses (GUMBEL 1958; WEISSMAN 1978).

Because we know the distribution of the top spacing, Δ_1 , we also know the distribution of beneficial effects when $i = 2$ and only one favorable mutant is available. It is

$$f(\Delta W|i = 2) = \frac{1}{E[\Delta_1]} e^{-\Delta W/E[\Delta_1]}. \quad (2)$$

In words, this distribution is exponential with mean $E[\Delta_1]$, as first noted by GILLESPIE (1991). We thus know the distribution of beneficial fitness effects among new mutations if adaptation always involves moving from the second-best ($i = 2$) to the best ($j = 1$) available allele.

But what if the wild-type allele has rank $i = 3$ and *two* favorable mutants are possible? If the population were to jump to the second-best allele, we have $f(\Delta W|i = 3, j = 2) = (2/E[\Delta_1]) \exp(-2\Delta W/E[\Delta_1])$; if the population were to jump to the best allele, we have (from the convolution), $f(\Delta W|i = 3, j = 1) = (2/E[\Delta_1]) [\exp(-\Delta W/E[\Delta_1]) - \exp(-2\Delta W/E[\Delta_1])]$. Substituting the previous two equations in (1), we find that the overall distribution of beneficial fitness effects is $f(\Delta W|i = 3) =$

$(1/E[\Delta_1]) \exp(-\Delta W/E[\Delta_1])$. This distribution is identical to that when starting at $i = 2$. Remarkably, this independence from fitness rank is a general result. This is proved in the next section where the moment-generating function (mgf) of $f(\Delta W|i)$ is derived.

The general result: To find the mgf of $f(\Delta W|i)$ we first find the mgf for ΔW conditional on mutating to a favorable allele of rank j (where $j < i$ as the mutation is beneficial). Because $\Delta W|i, j = \Delta_j + \Delta_{j+1} + \dots + \Delta_{i-1}$ and each Δ_n is independent, the mgf for $\Delta W|i, j$ equals the product of the mgf's for the individual Δ_n . But the Δ_n 's are exponentially distributed with means $E[\Delta_k] = E[\Delta_1]/k$. The conditional mgf is thus

$$M(t)|i, j = \left[\frac{1}{1 - E[\Delta_1]t/j} \right] \left[\frac{1}{1 - E[\Delta_1]t/(j+1)} \right] \cdots \left[\frac{1}{1 - E[\Delta_1]t/(i-1)} \right]. \quad (3)$$

The overall distribution of fitness effects is a mixture distribution, *i.e.*, one weighted by the probability of mutating to each favorable allele (Equation 1). Because the mgf of a mixture distribution equals the weighted average of the conditional mgf's (CHATFIELD and THEOBALD 1973) and the probability of mutating to a particular favorable allele is uniform over the integers $1 \leq j \leq i - 1$, we have

$$\begin{aligned} M(t)|i &= \sum_{j=1}^{i-1} M(t)|i, j \Pr\{j < i\} \\ &= \frac{1}{i-1} \sum_{j=1}^{i-1} \left(\prod_{k=j}^{i-1} \frac{1}{1 - E[\Delta_1]t/k} \right) \\ &= \frac{1}{1 - E[\Delta_1]t}. \end{aligned} \quad (4)$$

This is the mgf for an exponential distribution with mean $E[\Delta_1]$ and is independent of i . *Thus the distribution of fitness effects among beneficial mutations is exponential with mean $E[\Delta_1]$ independent of the fitness rank of the wild-type allele.* The spectrum of mutational fitness effects available to adaptation is, in other words, invariant: although the value of $E[\Delta_1]$ might vary from gene to gene, the expected distribution of beneficial effects at a given locus is independent of the fitness rank of the current wild type, where we assume only that wild-type fitness rank is high enough to use extreme value theory.

Figure 2 shows the results of exact computer simulations that test the accuracy of the above asymptotic the-

ory. A gene of length $L = 1000$ bp was simulated. For each distribution of allelic fitnesses, the fitnesses of $m = 3000$ mutant alleles plus one wild type were randomly drawn from the distribution of allelic fitnesses, ranked in fitness, and the difference in fitness between the wild type and a randomly chosen beneficial mutation was

recorded. Allelic fitnesses were assumed to be exponential, gamma, or half-normal (see Figure 2 legend for parameter values). Simulations were begun with wild-type fitness ranks of $i = 2, 10,$ or 25 (these different ranks translate into considerable differences in starting wild-type fitnesses given the distributions of allelic fitnesses used). Ten thousand replicates were performed for each set of conditions. Figure 2 shows that the theory nicely predicts the distribution of fitness effects among beneficial mutations regardless of the underlying distribution of allelic fitnesses. More important, the distribution of fitness effects among beneficial mutations is approximately invariant over a wide range of i , including those where it was unclear whether extreme value theory would hold (*e.g.*, $i = 25$), although some deviations appear at large i in the half-normal case. Because these deviations grow as i increases, it would seem unwise to extrapolate our results to much larger i (see Figure 2 legend).

It might seem that our results may simply reflect the memoryless property of exponential distributions. Many familiar distributions have an “exponential tail” in the sense that, as x gets large, $(1 - F(x + y))/(1 - F(x)) \rightarrow \exp(-cy)$, where c is a constant; in words, the probability of an increase of size y falls off exponentially and is independent of the precise “starting point,” x . The memoryless property of exponential tails cannot, however, fully explain our results. Many distributions of the Gumbel type do *not* show such tail behavior, *e.g.*, normal or lognormal distributions and those that are bounded on the right. Nonetheless, these distributions *still* have independent exponential extreme spacings and *still* give rise to an exponential distribution of fitness effects among new beneficial mutations. Our results hold asymptotically for all distributions in the domain of attraction of the Gumbel extreme value distribution whether or not they have “exponential tails” (see the APPENDIX).

Robustness of results: The above results are asymptotically independent of the shape of the distribution of allelic fitnesses (so long as it is of the Gumbel type) and i (so long as it is small). The above results are also robust to strong selection; indeed no weak selection approximations have been made.

Two assumptions, however, are potentially important. First, I assumed that each of the m mutations appears

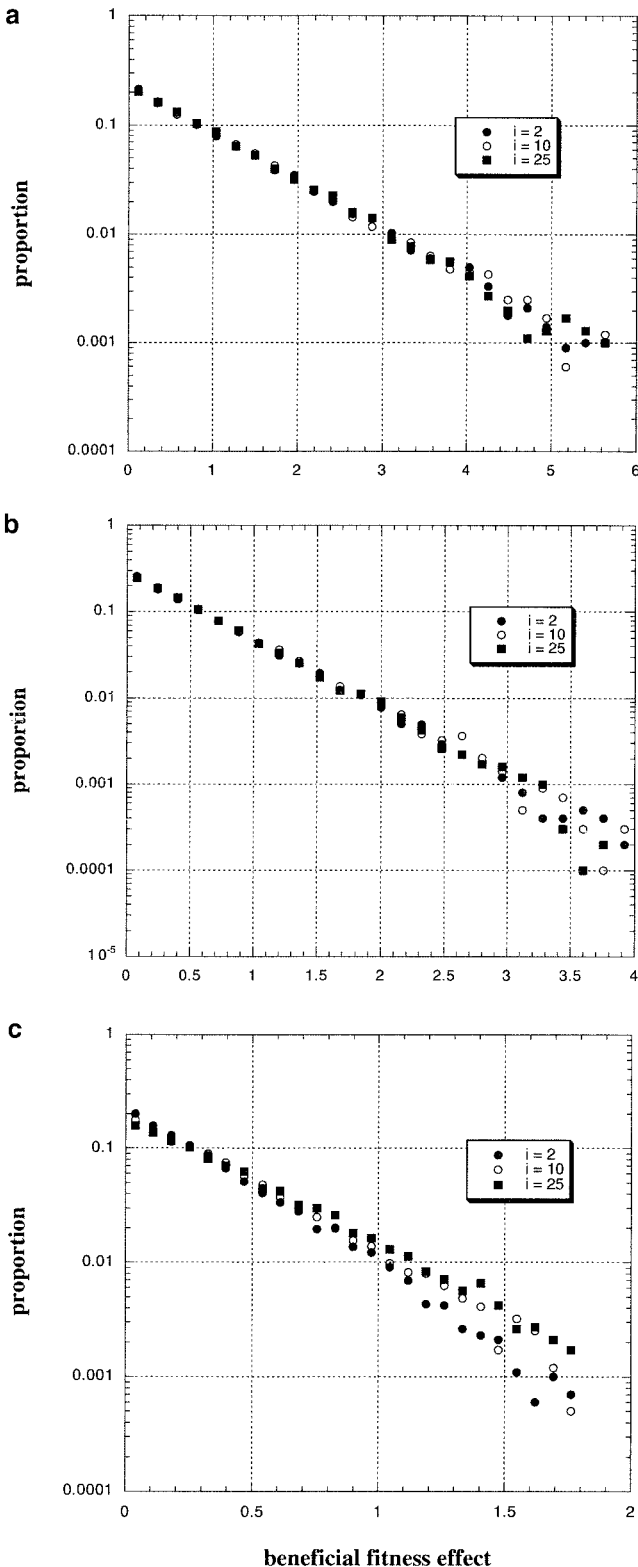


FIGURE 2.—Distribution of fitness effects among beneficial mutations: results of exact computer simulations. Three cases are shown: those in which the distributions of absolute allelic fitnesses were (a) exponential with mean = 1, (b) gamma with shape parameter of 2 and scale parameter of 2 (mean = 1), and (c) half-normal with mean = 1. The distribution of beneficial effects is exponential (straight line on a semilog plot) and essentially identical despite different starting wild-type fitness ranks ($i = 2, 10,$ and 25). This invariance holds even at larger i (*e.g.*, $i = 50$) in the exponential and gamma cases, although deviations occur in the half-normal case.

with equal frequency. It is easy to show that unequal mutation rates do not affect the above findings so long as all alleles are equally likely to have a given fitness rank; *i.e.*, mutationally common alleles are no more or less likely to have a given fitness rank than are mutationally rare alleles. Formally, the chance that the next mutation has rank j is $\Pr\{j\} = \sum_{k=1}^m \Pr\{\text{rank } j|k\} \Pr\{k\}$, where $\Pr\{k\}$ is the probability that the next mutation is to a particular allele k of the m possible. So long as all alleles are equally likely to have a given fitness rank, it is trivially true that $\Pr\{j\} = (1/m) \sum_{k=1}^m \Pr\{k\} = 1/m$, whether or not $\Pr\{k\}$ is the same for all m alleles. Considering only that subset of mutations that are beneficial, an analogous argument shows that $\Pr\{j|j < i\} = 1/(i-1)$, as in Equation 4. Thus Equation 4—and our key conclusion—remains correct whether or not all alleles are mutated to with equal frequencies. This was confirmed in computer simulations (not shown).

Second, following GILLESPIE (1983, 1984, 1991), I assumed that the distribution of allelic fitnesses is well behaved: it is a simple monotonically decreasing or unimodal distribution (*e.g.*, exponential, gamma, half-normal). But the actual distribution of allelic fitnesses at a locus might be a complicated mixture of several underlying distributions. To test the robustness of the analytic results, I used computer simulations to find the distribution of beneficial fitness effects when sampling from various “ugly” mixture distributions of allelic fitnesses. Figure 3 shows two such mixture distributions. One is a mixture of two underlying distributions and the other is a mixture of four underlying distributions. (In both cases, normal distributions contributed to the mixture distribution as the normal represents a near worst-case scenario, which converges to the extreme value distribution very slowly (GUMBEL 1958).) Figure 4 shows that the distribution of fitness effects among beneficial mutations remains roughly exponential in both cases. The distribution of fitness effects is also reasonably insensitive to starting i , which again ranged from $i = 2$ to $i = 25$. Figure 4 also shows, however, that as the tail of the mixture distribution grows lumpier, the distribution of beneficial effects becomes less well behaved. Our analytic results are therefore reasonably, but not indefinitely, robust to mixture distributions of allelic fitnesses.

DISCUSSION

Following GILLESPIE (1983, 1984, 1991), I have assumed that allelic fitnesses are drawn from some (unknown) probability distribution and that the present wild-type allele, while no longer the fittest sequence, is near the top in fitness. Under these assumptions, I have shown that the distribution of fitness effects (ΔW) among new beneficial mutations is exponential. More surprisingly, the distribution of beneficial effects shows an invariance property: it remains the same regardless of the fitness rank (and thus fitness) of the wild-type

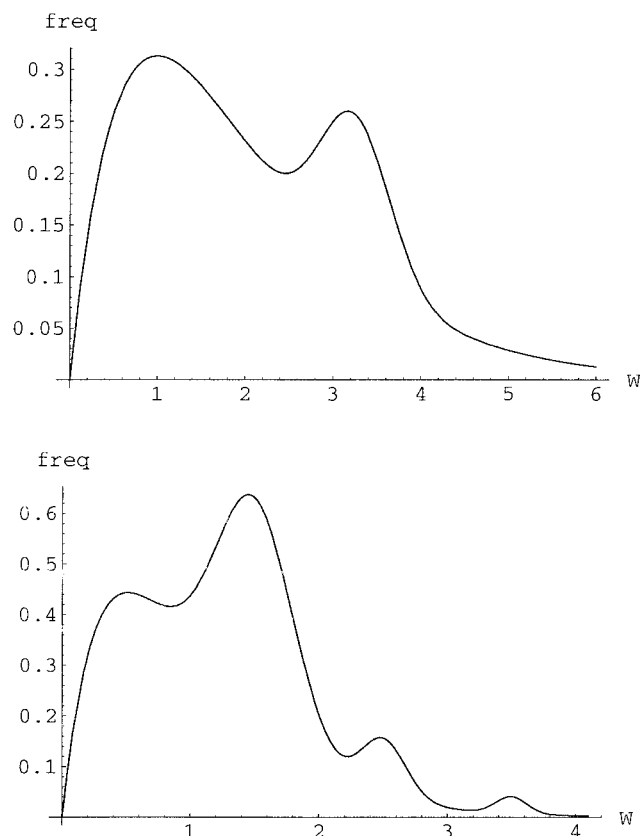


FIGURE 3.—Mixture distributions of allelic fitnesses. Top, allelic fitnesses are a mixture of two distributions (one gamma and the other normal). Bottom, allelic fitnesses are a mixture of four distributions (one gamma and three normal).

allele. Natural selection will, therefore, choose from the same spectrum of mutational effects whether adaptation starts from the second-best possible allele ($i = 2$) or one that is considerably worse (*e.g.*, $i = 10$).

These results depend on robust limit theorems from extreme value theory and so are quite general. They are independent of the distribution of allelic fitness (so long as it is of the Gumbel type), starting wild-type fitness (so long as it is high), strength of selection, and heterogeneity in mutation rates across sites. Although our results rest on asymptotic theory and so must be viewed as approximations (especially as different parent distributions converge on the extreme value distribution at different rates), computer simulations suggest that they are good approximations. Our results also hold in both sexual and asexual species and recombining and nonrecombining chromosome regions. There would seem to be good reason, then, for thinking that the distribution of beneficial fitness effects among new mutations at a locus might be generally approximately exponential and invariant. (There could, of course, be exceptions. One predicted by the present theory involves any locus at which the wild type is of very low fitness. Extreme value theory does not hold here. Another is where the

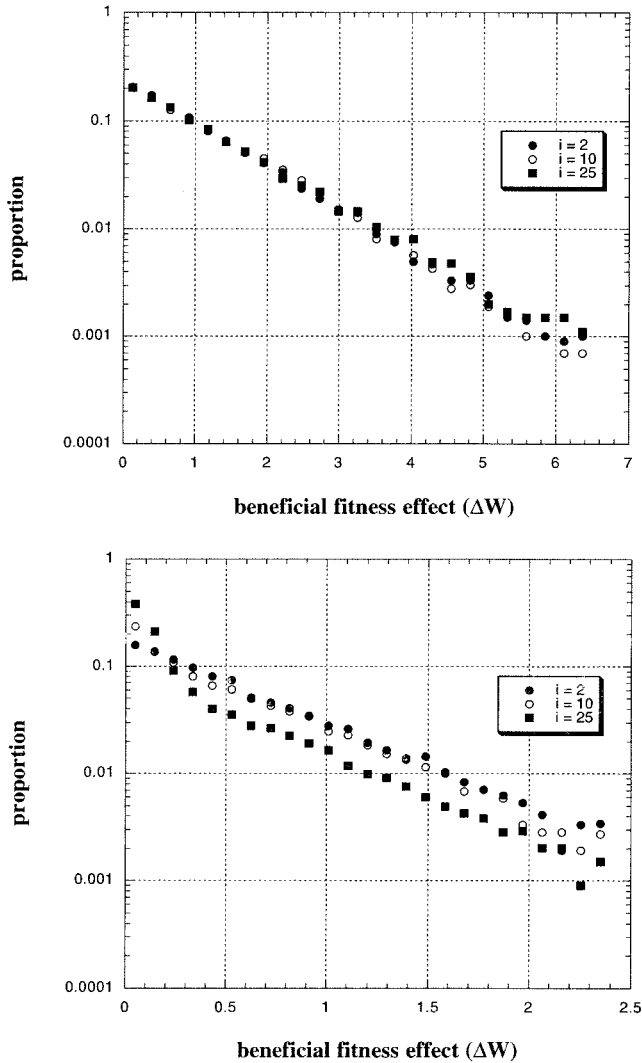


FIGURE 4.—The distribution of fitness effects among beneficial mutations when sampling from the mixture distributions of allelic fitnesses shown in Figure 3 (computer simulations; semilog plot). The top corresponds to a mixture of two distributions; the bottom corresponds to a mixture of four distributions.

distribution of allelic fitnesses has a very lumpy tail; see the above simulations.)

Though counterintuitive, the invariance property among beneficial mutations can be explained heuristically. If adaptation starts from a high-quality wild-type allele ($i = 2$), the jump to the best allele usually involves medium-sized fitness increases (Δ_1 ; see Figure 1). But if adaptation starts from a lower-quality allele ($i = 3$), jumps to better alleles involve some fitness increases that are usually smaller than before (Δ_2) and an equal number that are usually larger than before ($\Delta_1 + \Delta_2$). On average these balance and the mean fitness increase is unchanged. This argument generalizes for any starting wild-type fitness rank i , so long as it is small.

It is important to note that our results concern fitness increases, not selection coefficients. Selection coeffi-

cients are fitness increases normalized by wild-type fitness: $s = \Delta W/W_+$, where W_+ is the fitness of the wild-type allele. Because for any given i , ΔW and W_+ are both random variables, it is easy to show that selection coefficients do not enjoy the above invariance property. Instead the mean selection coefficient among beneficial mutations is $E[s] = E[\Delta_1/(W_1 - \Delta_1 - \dots - \Delta_{i-1})]$, which shrinks slightly with smaller i . Numerical work shows, however, that the distribution of s remains roughly exponential over small i (not shown; see also ROZEN *et al.* 2002).

The above theory, when combined with previous work, allows us to back calculate the mean fitness effect of *new* beneficial mutations from the mean fitness effect of *fixed* beneficial mutations, which are much more easily assayed in microbial experimental evolution work. Because large beneficial mutations have a greater chance of going to fixation than do small ones, the mean fitness increase among fixed beneficial mutations will obviously exceed (or at least equal) that among new beneficial mutations. ORR (2002) showed that, under the same assumptions as made here, the mean increase in fitness among fixed beneficial mutations in sexuals is $E[\Delta W_{\text{fixed}}] = 2(i-1)E[\Delta_1]/i$. This quantity ranges between $E[\Delta_1]$ and $2E[\Delta_1]$. Because the present theory shows that $E[\Delta_1]$ asymptotically equals the mean fitness effect of new beneficial mutations, it immediately follows that

$$\frac{E[\Delta W_{\text{fixed}}]}{2} \leq E[\Delta W_{\text{new}}] \leq E[\Delta W_{\text{fixed}}]; \quad (5)$$

i.e., the mean effect of new beneficial mutations is bounded between one-half and one times the mean effect of fixed beneficial mutations. Although this back calculation assumes that beneficial mutations enjoy independent fates, many asexual microbes having small genomes can be made to evolve under experimental conditions of effective sexuality (*e.g.*, sufficiently small population sizes that beneficial mutations arise one at a time). It thus appears that a notoriously elusive quantity—the mean fitness effect of new beneficial mutations—can be estimated in a way that is, at least in principle, straightforward.

Although theoretical population genetics has historically focused on neutral and deleterious mutations, recent theory has turned to adaptation (GERRISH and LENSKI 1998; HARTL and TAUBES 1998; ORR 1998, 2000, 2002, 2003; GERRISH 2001). This body of theory now lets us describe how a uniform rate of mutation to various mutant sequences gets transformed under fairly broad conditions into an exponential distribution of beneficial fitness effects of mean $E[\Delta W_{\text{new}}] = E[\Delta_1]$. In sexuals this distribution then gets transformed by probabilities of fixation into one of mean $E[\Delta W_{\text{fixed}}] = 2(i-1)E[\Delta_1]/i$ (ORR 2002). This distribution, which characterizes a single step in adaptation, in turn gets transformed during the stepwise approach to a fixed optimum into a

roughly exponential distribution of fitness effects (ORR 1998, 2002; the former article considered phenotypic effects and the latter selection coefficients; in both cases, however, ΔW is also roughly exponential). Thus both the distribution of beneficial effects among new mutations *and* the distribution of effects among the mutations ultimately fixed should be roughly exponential, at least when adaptation uses new mutations and approaches a constant optimum. It will obviously be of some importance to determine if similar patterns characterize adaptation when evolution proceeds from the standing genetic variation and/or approaches a moving optimum.

I thank N. Barton, A. Betancourt, P. Gerrish, J. Gillespie, J. P. Masly, D. Presgraves, M. Turelli, and two anonymous reviewers for helpful discussions or comments. I especially thank L. de Haan, I. Weissman, and D. Zelterman for helpful discussions of extreme value theory. This work was supported by National Institutes of Health grant 2R01 G51932-06A1 and by The David and Lucile Packard Foundation.

LITERATURE CITED

- ATWOOD, K. C., L. K. SCHNIEDER and F. J. RYAN, 1951 Selective mechanisms in bacteria. Cold Spring Harbor Symp. Quant. Biol. **16**: 345–355.
- BULL, J., M. BADGETT and H. WICHMAN, 2000 Big-benefit mutations in a bacteriophage inhibited with heat. Mol. Biol. Evol. **17**: 942–950.
- CHATFIELD, C., and C. M. THEOBALD, 1973 Mixtures and random sums. Statistician **22**: 281–287.
- DE HAAN, L., 1970 *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*. Tract 32, Amsterdam Mathematical Centre, Amsterdam.
- GERRISH, P., 2001 The rhythm of microbial adaptation. Nature **413**: 299–302.
- GERRISH, P. J., and R. E. LENSKI, 1998 The fate of competing beneficial mutations in an asexual population. Genetica **102/103**: 127–144.
- GILLESPIE, J. H., 1983 A simple stochastic gene substitution model. Theor. Popul. Biol. **23**: 202–215.
- GILLESPIE, J. H., 1984 Molecular evolution over the mutational landscape. Evolution **38**: 1116–1129.
- GILLESPIE, J. H., 1991 *The Causes of Molecular Evolution*. Oxford University Press, Oxford.
- GUMBEL, E. J., 1958 *Statistics of Extremes*. Columbia University Press, New York.
- HALDANE, J. B. S., 1924 A mathematical theory of natural and artificial selection, part I. Trans. Camb. Philos. Soc. **23**: 19–41.
- HALDANE, J. B. S., 1927 A mathematical theory of natural and artificial selection, part V: selection and mutation. Proc. Camb. Philos. Soc. **28**: 838–844.
- HARTL, D. L., and C. H. TAUBES, 1998 Towards a theory of evolutionary adaptation. Genetica **102/103**: 525–533.
- HOLDER, K., and J. BULL, 2001 Profiles of adaptation in two similar viruses. Genetics **159**: 1393–1404.
- IMHOF, M., and C. SCHLOTTERER, 2001 Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. Proc. Natl. Acad. Sci. USA **98**: 1113–1117.
- KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.
- LEADBETTER, M. R., G. LINDGREN and H. ROOTZEN, 1980 *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York.
- LENSKI, R. E., and M. TRAVISANO, 1994 Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc. Natl. Acad. Sci. USA **91**: 6808–6814.
- MAYNARD SMITH, J., 1970 Natural selection and the concept of a protein space. Nature **225**: 563–564.
- ORR, H. A., 1998 The population genetics of adaptation: the distribu-

tion of factors fixed during adaptive evolution. Evolution **52**: 935–949.

- ORR, H. A., 2000 Adaptation and the cost of complexity. Evolution **54**: 13–20.
- ORR, H. A., 2002 The population genetics of adaptation: the adaptation of DNA sequences. Evolution **56**: 1317–1330.
- ORR, H. A., 2003 A minimum on the mean number of steps taken in adaptive walks. J. Theor. Biol. **220**: 241–247.
- OTTO, S. P., and C. D. JONES, 2000 Detecting the undetected: estimating the total number of loci underlying a quantitative trait. Genetics **156**: 2093–2107.
- ROZEN, D. E., J. A. G. M. DE VISSER and P. J. GERRISH, 2002 Fitness effects of adaptations in microbial populations. Curr. Biol. **12**: 1040–1045.
- WAHL, L. M., and D. C. KRAKAUER, 2000 Models of experimental evolution: the role of genetic chance and selective necessity. Genetics **156**: 1437–1448.
- WEISSMAN, I., 1978 Estimation of parameters and large quantiles based on the k th largest observations. J. Am. Stat. Assoc. **73**: 812–815.
- WICHMAN, H. A., M. R. BADGETT, L. A. SCOTT, C. M. BOULIANNE and J. J. BULL, 1999 Different trajectories of parallel evolution during viral adaptation. Science **285**: 422–424.

Communicating editor: J. B. WALSH

APPENDIX

Most “ordinary” distributions belong to the Gumbel type (also known as type III). Here I briefly review the conditions for a distribution to belong to this type. My discussion is based loosely on that of LEADBETTER *et al.* (1980) and DE HAAN (1970).

Consider a parent distribution with probability density function (pdf) $f(x)$ and cumulative distribution function (cdf) $F(x)$. If we randomly draw a very large number, n , of values from this distribution, record the maximum value, and repeat this process many times, we will find that the distribution of the maximum (or of a linear function of the maximum) often tends to a limiting distribution, the so-called extreme value distribution. In reality, there are three extreme value distributions. “Ordinary” distributions like the normal, lognormal, exponential, gamma, etc., are in the domain of attraction of the Gumbel extreme value distribution; this is often casually referred to as *the* extreme value distribution. The cdf of the Gumbel distribution is $\Lambda(y) = \exp(-\exp(-y))$, where y is a linear function of the original random variable x . Many (though not all) bounded distributions are in the domain of attraction of another extreme value distribution, while exotic distributions like the Cauchy (whose moments are undefined) are in the domain of attraction of a third extreme value distribution. All of these extreme value distributions hold asymptotically as $n \rightarrow \infty$.

Parent distributions in the domain of attraction of the Gumbel distribution may have unbounded or bounded tails. The rightmost endpoint of the parent distribution is denoted x_F , where $x_F \leq \infty$ and $f(x) = 0$ for $x > x_F$. If $f(x)$ has a negative derivative over an interval (x_0, x_F) , a sufficient condition for $f(x)$ to be in the domain of attraction of the Gumbel extreme value distribution is that

$$\lim_{t \rightarrow x_F} \frac{f'(t)(1 - F(t))}{[f(t)]^2} = -1. \quad (\text{A1})$$

It is easy to show that familiar distributions like the exponential, gamma, normal, lognormal, etc., fulfill this condition.

The *necessary* and sufficient condition for $f(x)$ to be in the domain of attraction of the Gumbel extreme value distribution has also been found. It is

$$\lim_{t \rightarrow x_F} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-x}, \quad (\text{A2})$$

where $g(t)$ is a strictly positive function. It is important to note that this condition is not identical to having an “exponential tail” in the usual sense of $(1 - F(t + x)) / (1 - F(t)) \rightarrow \exp(-x)$ for large t . The reason is that

$g(t)$ need not be a constant. Indeed in the case of the normal and lognormal distributions, as well as in the case of bounded distributions, $g(t)$ is not a constant. Nonetheless, these distributions are in the domain of attraction of the Gumbel extreme value distribution.

If the maximum of a distribution converges to a particular extreme value distribution, the second and third, etc., largest-order statistics will converge to an asymptotic distribution of related functional form; *i.e.*, these order statistics belong to the same type as the maximum. WEISSMAN (1978) showed that all parent distributions in the domain of attraction of the Gumbel extreme value distribution have spacings between extreme order statistics that are asymptotically independent exponential random variables that behave as described in the text.