Genetica **102/103:** 349–358, 1998.
© 1998 *Kluwer Academic Publishers. Printed in the Netherlands.*

Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*

Santiago F. Elena, Lynette Ekunwe, Neerja Hajela, Shenandoah A. Oden & Richard E. Lenski*
*Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824, USA; * Author for correspondence (Fax: 517-353-3955; e-mail: lenski@pilot.msu.edu)*

Key words: deleterious mutations, *Escherichia coli*, fitness, insertion mutations, transposition

Abstract

Very little is known about the distribution of mutational effects on organismal fitness, despite the fundamental importance of this information for the study of evolution. This lack of information reflects the fact that it is generally difficult to quantify the dynamic effects of mutation and natural selection using only static distributions of allele frequencies. In this study, we took a direct approach to measuring the effects of mutations on fitness. We used transposon-mutagenesis to create 226 mutant clones of *Escherichia coli*. Each mutant clone carried a single random insertion of a derivative of Tn*10*. All 226 mutants were independently derived from the same progenitor clone, which was obtained from a population that had evolved in a constant laboratory environment for 10,000 generations. We then performed competition experiments to measure the effect of each mutation on fitness relative to a common competitor. At least 80% of the mutations had a significant negative effect on fitness, whereas none of the mutations had a significant positive effect. The mutations reduced fitness by about 3%, on average, but the distribution of fitness effects was highly skewed and had a long, flat tail. A compound distribution, which includes both gamma and uniform components, provided an excellent fit to the observed fitness values.

Introduction

Mutations influence the rate of almost every evolutionary process, from the incorporation of novel beneficial alleles by natural selection (Haldane, 1927) to the substitution of neutral alleles by random drift (Kimura, 1983) and the spread of deleterious mutations by Muller's ratchet (Muller, 1964). The distribution of mutational effects on phenotypic traits, and ultimately on fitness, is therefore of fundamental importance for predicting evolutionary dynamics (Crow & Kimura, 1970; Lande, 1975, 1983; Turelli, 1984; Hill & Rasbash, 1986; Barton & Turelli, 1987). Yet, there exists surprisingly little quantitative information on the distribution of mutational effects.

Some studies have estimated the total genomic mutation rate for various organisms, in some cases including all mutations (Drake, 1991), while in other cases including only those mutations that are deleterious (Mukai, 1964; Mukai et al., 1972; Houle et al., 1992; Kibota & Lynch, 1996). Many other studies have estimated the contribution of mutation to standing genetic variation but without directly estimating either mutation rates or mutational effects (e.g., Clark, Wang & Hulleberg, 1995a; for a review, see Houle, Morikawa & Lynch, 1996). A few ambitious studies have even sought to measure the distribution of mutational effects. However, these studies have focused either on phenotypic traits, such as bristle number and enzyme activity in *Drosophila melanogaster*, that are of unclear adaptive significance, or on viability, which represents only one component of fitness (Mackay, Lyman & Jackson, 1992; López & López-Fanjul, 1993; Clark, Wang & Hulleberg, 1995b; Lyman et al., 1996).

Bacteria such as *E. coli* provide a powerful experimental system for addressing a wide range of evolutionary questions (see reviews by Dykhuizen & Hartl, 1983; Levin & Lenski, 1983; Dykhuizen & Dean, 1990; Lenski, 1992; Sniegowski & Lenski, 1995). They are easy to handle and propagate, which allows

intensive replication of experiments so that even subtle effects can be measured. Importantly, one can directly estimate the overall relative fitness of genotypes by placing them in competition and measuring the ratio of their population growth rates over many generations. More than 1500 competition experiments were performed in this study alone, with most lasting for 40 generations. Also, a variety of molecular techniques are available for the construction of genotypes of interest. In the present study, we used mutagenesis with mini-Tn*10* transposon derivatives to produce more than 200 genotypes, all derived from the same progenitor but each one containing a different single random insertion mutation. These mutations are stable, they are easily detected by the presence of a marker encoded by the transposon, and their insertion sites can be mapped using Southern blots or sequence analysis.

Materials and methods

Bacterial genotypes

All of the mutants generated in this study were derived from a single progenitor clone of *E. coli* B, which is designated REL4548. This clone was sampled from an experimental population, designated $Ara^{-}1$, that had been serially propagated for 10,000 generations in a minimal medium supplemented with glucose (Lenski et al., 1991; Lenski & Travisano, 1994). This particular population maintained the low genomic mutation rate characteristic of most *E. coli* strains, whereas some other replicate populations had evolved mutator phenotypes (Sniegowski, Gerrish & Lenski, 1997). During this period, the mean fitness of the population increased by about 50% relative to its ancestor. However, the rate of fitness increase slowed dramatically, indicating that the population had approached a quasi-equilibrium with respect to selection. A slow rate of improvement is presumably similar to the state of most natural populations, except those that have recently encountered novel environmental conditions.

This progenitor clone is unable to metabolize the sugar arabinose. We obtained a spontaneous Ara^{+} mutant of another clone from the same population by plating millions of cells onto minimal agar that contained arabinose as a sole carbon source. This Ara^{+} mutant (SFE/I+1) was used as the common competitor in the experiments to measure the relative fitness of each mutant genotype, as described below. Ara^{-} and Ara^{+} cells are readily distinguished by their production of red and white colonies, respectively, on tetrazolium arabinose (TA) indicator agar (Levin, Stewart & Chao, 1977; Lenski, 1988).

Transposons and delivery vectors

Mutations were generated by random insertion of mini-Tn*10* derivatives, following the methods of Kleckner, Bender and Gottesman (1991). This approach has several important advantages. Secondary transpositions cannot occur because the transposase is not expressed by the mini-Tn*10* construct, but rather it is encoded by the delivery vector. Also, the transposase gene has two ATS (altered target specificity) mutations, such that the mini-Tn*10* elements are inserted more or less at random into the bacterial chromosome. Finally, mini-Tn*10* derivatives carry antibiotic-resistance markers that allow easy detection of mutated cells on selective media.

We used three different mini-Tn*10* derivatives, each of which encoded one of the following markers: chloramphenicol resistance (Cam^{R}), kanamycin resistance (Kan^{R}), and tetracycline resistance (Tet^{R}). Resistance to tetracycline is inducible, whereas the other two resistance genes are expressed constitutively. In the absence of an inducer, expression of the resistance protein is very tightly repressed (Nguyen et al., 1989), therefore minimizing the possibility that the resistance marker itself would have a systematic effect on the mutant's fitness. For this reason, we performed the majority of our experiments using the Tet^{R} marker.

The delivery vector for the mini-Tn*10* derivatives is a defective Lambda phage, λNK, which has two important features (Kleckner, Bender & Gottesman, 1991). First, the phage has been crippled by nonsense mutations in all of the genes required for its autonomous replication and lytic growth; the phage, therefore, requires for its replication a special host, *E. coli* C600 $RK^{-}MK^{+}$. Second, the phage cannot lysogenize cells because it has been rendered defective in its integration functions. Plasmids (pNK derivatives) that carried the three mini-transposons were used to make probes for molecular confirmation of insertion mutations. Plasmids and phages, as well as the special host, were obtained from the American Type Culture Collection.

Mutagenesis and selection of mutants

The protocol for generating insertion mutations was that of Kleckner, Bender and Gottesman (1991). Tar-

get cells (REL4548) were grown overnight in tryptone-thiamine-maltose medium to yield $\sim 10^{10}$ cells, pelleted, and resuspended in 1 ml of LB. Then, 0.1 ml of the concentrated target cells were infected with the λNK phage at a multiplicity of infection of 0.1 phage per cell. After allowing 15 min for phage adsorption to the target cells, any remaining free phage were removed by washing twice with 5 ml of LB supplemented with 50 mm sodium citrate. The infected cells were then diluted in 5 ml of LB plus citrate and incubated at 37 °C for 1 h. At this time, 0.1 ml of the infected cell mixture was spread on TA agar plates supplemented with an antibiotic that selected for cells carrying the mini-Tn*10* derivative. TA agar is a rich nutritive medium that permits the growth even of auxotrophic mutants; hence, it should allow the detection even of those mutants that would be severely handicapped in a more restrictive minimal medium.

A pilot experiment indicated that no additional mutant colonies appeared on the antibiotic-containing selective plates after 72 h of incubation at 37 °C. In order to avoid possible biases against mutant genotypes that grow more slowly, we obtained a stratified sample of mutants, based on the time course of their colony appearance. However, as we will show in the Results section, there was no significant association between the time of colony appearance and relative fitness. We obtained 176 mutants that carried a Tet^Rmini-Tn*10* insertion and an additional 25 each that carried $Cam^R$and Kan^R insertions.

Molecular confirmation of random insertions

As noted above, the transposase expressed from the delivery vector was genetically modified so that the mini-Tn*10* derivatives should be inserted at random locations in the bacterial chromosome (Kleckner, Bender & Gottesman, 1991). We checked this expectation, without actually mapping each insertion, by performing Southern-blot analyses on 60 of the mutants. First, we made probes from the pNK plasmids by extracting a fragment of the relevant resistance genes by means of double digestions. In the case of pNK(Cam), digestion with *Bcl*I and *Eco*RI generated a fragment of $\sim$ 700 bp. For pNK(Kan), digestion with *Ava*I and *Pvu*I yielded a $\sim$ 400 bp fragment. Digestion of pNK(Tet) with *Bgl*I and *Xba*I gave a fragment of $\sim$ 900 bp. These three probes were purified from 0.8% agarose gels using 0.45 μm filters (Millipore Ultrafree®-MC) and labeled using the random-priming digoxigenin method (Boehringer-Mannheim). The genomic DNA of each mutant clone was then purified using a genomic DNA purification kit (Quiagen). A total of 2.5 μg of genomic DNA was digested with *Sac*I and *Kpn*I and subjected to electrophoresis overnight at 1 V/cm in 0.5% agarose and 1 $\times$ TBE. These particular restriction enzymes were chosen because they do not cut anywhere inside the mini-Tn*10* constructs and thus were expected to yield single hybridization signals. DNA was then transferred to Biotrans (+) nylon membranes (ICN) using a vacuum blotting system (Pharmacia-Biotec) and UV-crosslinked to the membrane. Hybridization to the digoxigenin-labelled probe and subsequent detection of the colored signal were done according to the recommendations of the supplier (Boehringer-Mannheim).

Fitness estimation

We performed competition experiments to estimate the fitness of each mutant clone relative to the Ara^+ common competitor. As noted above, the Ara^+ marker allows genotypes to be readily distinguished by colony color on TA agar plates. Competition experiments were performed using the same culture conditions as used for the long-term evolution experiment from which the unmutated progenitor was sampled after 10,000 generations (Lenski et al., 1991; Lenski & Travisano, 1994). Populations were grown in a shaking incubator at 120 rpm and 37 °C in Davis minimal medium supplemented with 25 μg/ml glucose (DM25). Every day, 0.1 ml of culture was transferred into 9.9 ml of fresh DM25, which allowed $\sim$ 6.6 (= $\log_2 100$) generations of binary fission per day. Prior to each competition experiment, both the mutant carrying the mini-Tn*10* insert and the Ara^+ common competitor were grown separately for one day in DM25 to ensure that they were comparably acclimated to the competition environment. The two competitors were then mixed at a 1:1 volumetric ratio and diluted 100-fold into fresh DM25. This mixed culture was serially propagated by daily transfers into fresh DM, usually for 6 days ($\approx$40 generations). Initial and final samples of the mixed culture were spread onto TA agar plates to determine initial and final densities of both competitors. The fitness of a few mutants was sufficiently low relative to the common competitor that after 6 days the mutants comprised only a small fraction of the total population ($<$5%). In those cases, the competition experiments were rerun for either 4, 2, or 1 transfer cycles (depending upon the severity of the mutant's competitive disadvantage) in order to obtain more precise estimates of relative fitness.

We computed each competitor's net rate of population growth, or realized Malthusian parameter, as $M = \ln (N_t \times 100^t / N_0)$, where N_0 and N_t are the stationary-phase population densities at time zero and after t days, respectively. The fitness of one competitor relative to the other was then calculated simply as the ratio of their realized Malthusian parameters during head-to-head competition (Lenski et al., 1991). To standardize fitness values relative to the unmutated progenitor, we performed a total of 195 replicate competitions between the progenitor and the Ara^+ common competitor. The fitness of each mutant was then standardized relative to the mean value of the 195 fitness estimates for the unmutated progenitor. One mutant was simply unable to grow in DM25, and so its relative fitness was zero, by definition. For the other 225 mutants, we first performed three or five replicate competitions against the Ara^+ common competitor. To examine the possibility of type I and type II statistical errors, we then performed an additional 10 or 30 competitions for a subset of the mutants, as described in the Results section.

Statistical analyses

To determine whether any given mutation had a significant effect on fitness, we performed a bootstrap significance test to compare the N fitness values obtained for the mutant clone with the 195 fitness values obtained for the unmutated progenitor. We employed the bootstrap because it does not assume equal variances; the sampling variance for fitness tends to increase at low fitness values because the final sample typically contains fewer individuals of the inferior competitor. Specifically, for each mutant, we generated 10,000 bootstrap samples (with replacement) from the observed data, with each bootstrap sample including N and 195 fitness values for the mutant and progenitor clones, respectively. For each bootstrap sample, we then calculated the ratio of the mutant's fitness to that of the progenitor. We excluded the most extreme 2.5% of the bootstrap samples in each tail, and we judged the mutation to have had a significant effect on fitness if the value 1.0 was outside the resulting 95% confidence interval.

For the purpose of describing the population distribution of mutational effects on fitness, each mutant was treated as an independent observation. The significance values of the skewness (g_1) and kurtosis (g_2) statistics were tested as described by Sokal & Rohlf (1981).

Results

Physical characterization of mutant clones

Sixty of the mutant clones were examined by Southern-blot analyses. All of them yielded a single hybridization signal, demonstrating that each clone carried a mini-Tn*10* insertion and, moreover, that no secondary transpositions had occurred. None of the resulting DNA fragments that carried these insertions were the same physical size, with the possible exception of a few large fragments for which it was difficult to resolve their sizes with any accuracy (data not shown). The insertion mutations were, therefore, dispersed throughout the chromosome instead of being concentrated in a few insertion hot-spots.

Effect of different markers on fitness

Figure 1 shows the distributions of fitness values measured for the three sets of insertion mutations, which were generated using mini-Tn*10* constructs with three different antibiotic-resistance markers. The mean fitness for the 176 Tet^R clones was 0.9708, with a standard deviation of 0.0934 (Figure 1a). For the 25 Cam^R clones, the mean fitness was 0.9684, with a corresponding standard deviation of 0.0620 (Figure 1b). And the mean fitness of the 25 Kan^R clones was 0.9882, with a standard deviation of 0.0068 (Figure 1c). Notice that the much higher standard deviations for the Tet^R and Cam^R sets than for the Kan^R set are due to a handful of clones in the former two sets that had extremely low fitness. We performed two statistical tests to ascertain whether the three sets of mutants had significantly different fitness distributions. We used nonparametric tests, because the fitness values were not normally distributed (see below). We first performed a Kruskal-Wallis test, which uses rank-ordered values to assess differences in central location among the three sets of mutants (Sokal & Rohlf, 1981). No compelling effect of the resistance markers was found (H = 5.3067, 2 df, P = 0.0704). We then performed three Kolmogorov-Smirnov two-sample tests to compare the overall distribution of fitness values for each pair of sets (Sokal & Rohlf, 1981). None of the three pairwise tests approached statistical significance (all $P >$ 0.1). Thus, differences among clones in the particular antibiotic-resistance marker that they carried had little or no effect on their relative fitness values. Therefore, we combined the three sets of mutants for all of our subsequent analyses.

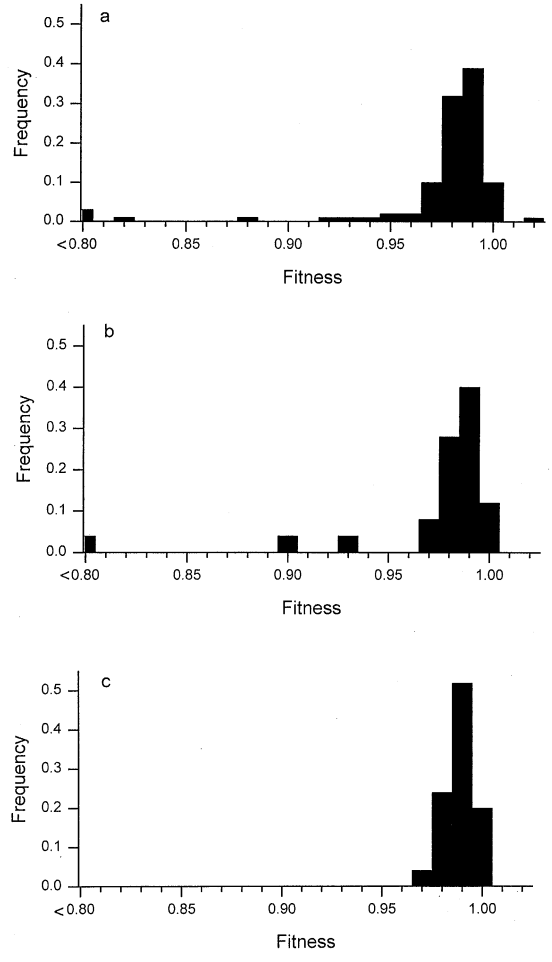

Figure 1. Distributions of relative fitness values measured for three sets of insertion mutations generated using mini-Tn*10* derivatives. (a) Distribution for 176 clones that carry the Tet^R marker. (b) Distribution for 25 clones that carry the Cam^R marker. (c) Distribution for 25 clones that carry the Kan^R marker. The three distributions are not significantly different from one another (see text).

Correlation between appearance time and fitness

Some of the colonies that gave rise to the mutant clones appeared after one day in the initial antibiotic-selection screen, whereas others took two or three days to form a visible colony. However, there was no significant correlation between a mutant's competitive fitness in DM25 and the day of its appearance in the initial screen (r = -0.0561, 224 df, P = 0.4013). Evidently, whether a mutant colony was detected early or late during the initial screen had little or no bearing on its competitive fitness in the experimental environment.

Effects of individual insertion mutations on fitness

Of the 226 clones that carried an insertion mutation, 169 yielded fitness values that were below 1.0 at the $P< 0.05$ significance level, based on the initial 3 to 5 replicate competition experiments that were performed for each clone. In striking contrast, only one clone showed a significant improvement in fitness relative to the unmutated progenitor.

We then sought to investigate further the possibility of type I and II statistical errors (Sokal & Rohlf, 1981). A type I error occurs when, by chance, a significant effect is claimed even though the null hypothesis is actually true (or when the actual effect goes in the opposite direction to the observed effect, which is judged significant, but in the wrong direction). A type II error occurs when the null hypothesis is actually false, but the statistical test lacks sufficient power to detect that effect. Concerning the 169 cases where mutations were judged to have significant deleterious effects on fitness, it is certainly possible that *a few* of these mutations were actually neutral but fell into the deleterious category by virtue of type I errors. Given that we performed 226 tests, and assuming for the sake of illustration that all of the mutations were truly neutral, we would expect about 6 ($\approx 0.05 \times 226 / 2$) mutations to have shown 'significant' deleterious effects by chance alone. Nonetheless, it is also evident that not more than a tiny fraction of the 169 cases of significant deleterious effects can be attributed to type I errors. In fact, for 113 mutant clones, the reduction in fitness was significant even at the $P< 0.001$ level, and it is unlikely that even one of these cases can be attributed to a type I error.

The one case in which a random insertion mutation was judged to have a significant positive effect is more interesting with respect to the possibility of a type I error. Let us assume, for the purpose of illustration, that 169 mutations were actually deleterious, with the other 57 mutations truly neutral. Given 57 neutral mutations, we would expect one or two ($\approx 0.05 \times 57 / 2$) of them to exhibit a 'significant' beneficial effect due to chance alone. To test this possibility of a false-positive more thoroughly, we performed an additional 30 competitions with the mutant clone that initially gave a statistically significant improvement in its fitness. When we did so, the beneficial effect was not supported and, in fact, this mutant exhibited a small (0.8%) but significant reduction in its fitness relative to the unmutated progenitor (P= 0.002).

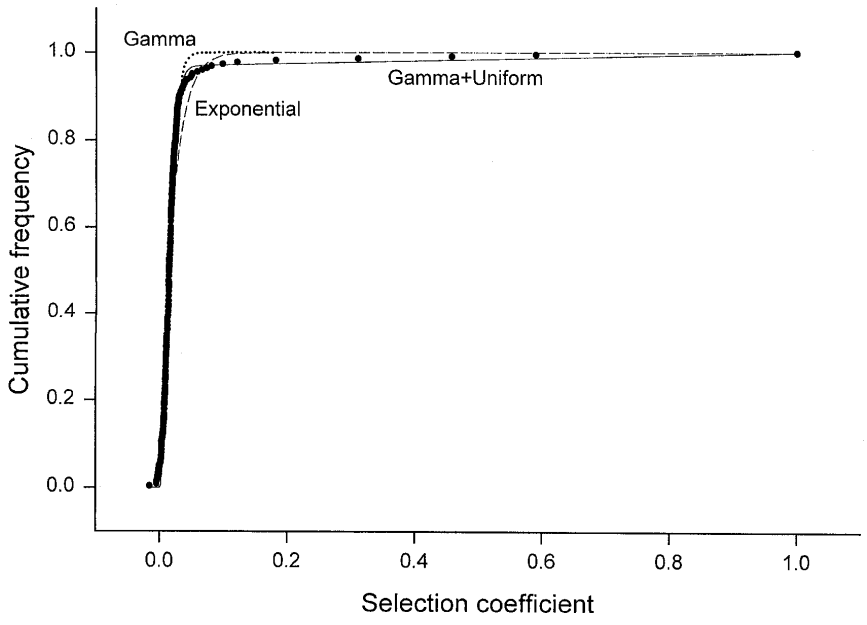

Figure 2. Cumulative frequency distribution for selection coefficients associated with insertion mutations. The filled circles indicate the observed distribution for 226 different mutations. The dashed curve shows the best fit of a negative exponential distribution to these data. The dotted curve shows the best of a gamma distribution to these data. The solid curve shows the best fit of a compound gamma plus uniform distribution to the data. See text for details and statistical comparisons.

To examine the possibility of type II errors, we chose 13 mutant clones whose fitness values were not significantly different from the unmutated progenitor, based on the initial set of competition experiments. We then ran an additional 10 competitions with each of these clones. In all 13 cases, the more extensive data indicated that the insertion mutation did indeed cause a significant fitness loss (in all cases between 1.2% and 2.2%) relative to the unmutated progenitor (all $P <$ 0.001). Evidently, most of the mutations that appeared to be neutral based on the initial set of competitions are, in fact, slightly deleterious.

In summary, at least 80% of the insertion mutations (169 + 1 + 13 out of 226 total) caused a significant reduction in competitive fitness. None of the mutations conferred a significant advantage. The remaining 20% may be selectively neutral, or their fitness effects may have been too small to detect given the statistical power of our experiments.

Distribution of fitness effects

The mean fitness over all 226 mutant clones was 0.9725 with a standard deviation of 0.0850. A Kolmogorov-Smirnov test for goodness of fit (Sokal & Rohlf, 1981) showed that the distribution of fitness effects deviates significantly from normality (D= 0.3924, Z = 5.8990, $P < 0.0001$). The distribution is highly skewed towards low fitness values (g_1= -8.7872, t_{225} = 54.2755, $P <$ 0.0001), and consequently the median fitness (0.9860) is well above the mean fitness. The distribution is also strongly leptokurtic (g_2= 87.4483, t_{225} = 271.2416, $P < 0.0001$), such that many values lie near the center and in the tail whereas relatively few have intermediate values.

Fit of observed distribution to alternative models

The fitness effects caused by insertion of the mini-Tn*10* elements are evidently not normally distributed, and so we sought to examine the fit of these data to alternative models that have been used in the literature to describe the distribution of mutational effects on fitness. We performed least-squares nonlinear regression using the Levenberg-Marquardt method (Press et al., 1990) to obtain the best fit of each model to the empirical data. We used several different initial parameter estimates to ensure convergence on the final estimates.

The simplest model that captures the basic feature that many more mutations have small effects than have large effects is the negative exponential distribution (Gregory, 1965; Mukai et al., 1972; Ohta, 1977). This

model has only a single parameter, $1/\overline{s}$, which equals the inverse of the average mutational effect on fitness. The dashed curve in Figure 2 shows the best fit of this model to the data, and it yields an estimate of $\overline{s}$ = 0.0197. This simple model provides a reasonably good fit to the data (R^2 = 0.909, 225 df, $P < 0.0001$). However, almost 10% of the observed variation in mutational effects is still unexplained.

An alternative model, which employs two parameters, is the gamma distribution (Kimura, 1979; Lynch & Gabriel, 1990; Keightley, 1994). Roughly speaking, α is a scaling factor, while β determines the shape of the fitness distribution, although in fact both parameters affect the mean, variance, and kurtosis of the distribution (Kimura, 1979). The average selection coefficient, $\overline{s}$, is equal to the ratio β / α. The dotted curve in Figure 2 shows the best fit of the gamma distribution to our data (β = 2.81, α = 173.79, and $\overline{s}$ = 0.0162). The gamma model requires an additional parameter, but a likelihood-ratio test (Sokal & Rohlf, 1981) indicates that the gamma distribution gives a much better fit to the data than did the negative exponential distribution (G = 79.6554, 1 df, $P < 0.0001$).

Although the gamma distribution provides a very good fit to the data, it does not adequately reflect the fitness distribution for those few mutations with large selection coefficients ($>>$ 3%). Keightley (1994, 1996) proposed that mutational effects may be better described by a compound model, in which two (or more) component distributions are mixed to yield the overall distribution. Given that the gamma distribution failed to capture the long, flat tail of the observed distribution of mutational effects in this study (Figure 2), we considered a compound model in which a fraction, p, of the mutants are drawn from a gamma distribution while the remaining fraction, 1 - p, are drawn from a completely uniform distribution ($0 \leq s \leq 1$). The solid curve in Figure 2 shows the best fit of this compound model to the empirical distribution (p = 0.9683, β = 3.03, and α = 194.24). Unlike the pure exponential and gamma models, this compound distribution reproduces the long, flat tail of the empirical distribution. The average selection coefficient of the compound model, $\overline{s}$, is equal to 0.0310 [= $p \times (\beta / \alpha) + (1 - p) \times 0.5$], which agrees much better with the observed value of 0.0275 than did either the exponential or gamma model. The compound model uses a third parameter, but even so a likelihood-ratio test indicates that its improved fit in comparison with the pure gamma model is highly significant (G = 75.8290, 1 df, $P < 0.0001$). This 3-parameter model leaves unexplained only 0.5% of the observed variation in mutational effects.

Discussion

The distribution of mutational effects on fitness is a critical component of many mathematical models of evolutionary dynamics (Crow & Kimura, 1970; Lande, 1975, 1983; Turelli, 1984; Hill & Rasbash, 1986; Barton & Turelli, 1987), but there is very little quantitative information on this distribution (Barton & Turelli, 1989; Mackay, Lyman & Jackson, 1992; López & López-Fanjul, 1993; Lyman et al., 1996). To address this deficit, we used transposon-mutagenesis to create 226 clones of *E. coli*, each carrying a different single mutation caused by the random insertion of a mini-Tn*10* derivative (Kleckner, Bender & Gottesman, 1991). These mutants were derived from the same progenitor clone, which came from a population that had evolved in a constant laboratory environment for 10,000 generations and was evidently near a selective equilibrium (Lenski & Travisano, 1994). Competition experiments indicated that at least 80% of the mutations had a significant negative effect on fitness, while none of the 226 mutations examined had a significant beneficial effect. The mutations reduced fitness by about 3%, on average, but the distribution of fitness effects was highly skewed and had a long, flat tail (Figure 1).

Several lines of evidence indicate that the observed fitness effects were primarily due to the insertion mutations *per se*, as opposed to the genetic markers used to screen for these mutations. First, the distributions of fitness effects were statistically indistinguishable for the three different antibiotic-resistance markers that we used, despite the fact that their physiological mechanisms of resistance are completely different. For example, resistance to chloramphenicol occurs by the enzymatic acetylation of the antibiotic, which renders it nontoxic to the cell, whereas resistance to tetracycline involves the active efflux of that compound from the cell. Second, tetracycline resistance is extremely tightly repressed in the absence of antibiotic (Nguyen et al., 1989). Hence, resistance would have been expressed phenotypically during the initial screening for mutations, but not during the competition experiments to estimate relative fitness. Third, there is tremendous variation among the individual mutations in their fitness values (Figure 1), which would not be expected if the deleterious effects were due to the markers. The

variation in fitness among mutant clones is highly significant even for those mutations with relatively mild effects. Recall the supplementary experiment in which we measured, with ten-fold replication, the fitness of 13 clones, all marked with Tet^R, that had small selection coefficients (ranging from 1.2% to 2.2%). Despite this very narrow range of fitness values, an analysis of variance indicates that the heterogeneity of mutational effects is highly significant (F = 13.936, 12 and 117 df, $P< 0.0001$).

We cannot absolutely exclude the possibility that some of the strains carrying insertion mutations might have additional mutations that occurred during their construction. However, the genomic mutation rate for most strains of *E. coli*, including the progenitor clone used in this study, is very low (Drake, 1991; Kibota & Lynch, 1996; Sniegowski, Gerrish & Lenski, 1997), so that it is unlikely that background mutations are an important factor in the observed distribution of fitness effects. The act of transposon insertion itself might sometimes cause secondary mutations; to the extent that such associated mutations occurred in this study, their effects on fitness are inseparable from, and included with, the effects of the insertions *per se*.

A compound distribution, with both gamma and uniform components, gave an excellent fit to the observed selection coefficients (Figure 2). The fit of this compound distribution was significantly better than those provided by pure uniform, exponential, and gamma distributions, which have been used by most theoreticians and experimentalists (Gregory, 1965; Mukai et al., 1972; Ohta, 1977; Kimura, 1979; Lynch & Gabriel, 1990; Charlesworth, Morgan & Charlesworth, 1993; Kondrashov, 1994; Lande, 1994, Butcher, 1995; Lynch, Conery & Bürger, 1995; but for a previous use of a different compound distribution, see Keightley, 1996). Although the uniform component represented only a small fraction of the compound distribution, it accounted for approximately half of the average selection coefficient; that is, the average selection coefficient was only about 1.5% excluding the long, flat tail of the distribution. Of course, in a population that is at mutation-selection balance, those relatively few but highly deleterious mutations will be found at correspondingly very low frequencies compared to less harmful mutations. At equilibrium, the contribution of a particular deleterious mutation to a population's genetic load depends on the underlying mutation rate but not on the magnitude of its deleterious effect (Haldane, 1937; Muller, 1950; Crow & Kimura, 1970). In contrast, the expected genetic variance at equilibrium very much depends on the distribution of mutational effects (Lande, 1975; Turelli, 1984; Keightley & Hill, 1988).

Chao et al. (1983) have shown that Tn*10* can cause beneficial mutations in experimental populations of *E. coli* and thereby accelerate their rate of adaptive evolution. However, the lack of any beneficial mutations among the 226 cases examined in this study is not surprising, for at least two reasons. First, the particular mutations detected by Chao et al. were not typical members of the set of all possible Tn*10*-induced mutations, but rather these mutations were from the subset that was enriched by natural selection acting in a very large population ($>10^{10}$ cells). In fact, when the number of cells that carried Tn*10* was too low ($<10^6$ or so), these transposon-generated beneficial mutations were not detected even over several tens of generations, implying a corresponding mutation rate less than 10^{-6}. Second, based on theoretical considerations, Gerrish and Lenski (1998) have estimated that only about one mutation in 10^6 is beneficial for the particular *E. coli* strain and environment used in this study. In fact, this estimate was for the ancestral strain that had not yet adapted to the experimental environment, whereas we used a strain that had already evolved in this environment for 10,000 generations and had, therefore, presumably 'used up' some of these beneficial mutations.

We can also compare our direct estimate of the average selection coefficient with an indirect estimate recently obtained by Kibota and Lynch (1996) from mutation-accumulation experiments with *E. coli*. Kibota and Lynch used as their progenitor a clone sampled from an earlier time-point in the same 10,000-generation experiment (Lenski et al., 1991; Lenski & Travisano, 1994), and they used a similar environment. Kibota and Lynch obtained mutated lines by repeatedly passing populations through single-cell bottlenecks. From the resulting trajectories for the mean and variance in fitness, they obtained an upper-bounded estimate of the average deleterious mutational effect, $\overline{s}_{max}$, that was equal to 0.012. This value is somewhat lower than our estimate of $\overline{s}$, which is 0.031. In fact, there are two compelling reasons to expect *a priori* that our approach would yield a higher estimate than the approach used by Kibota and Lynch. First, we examined only insertion mutations, whereas their data would reflect all spontaneous mutations, including point mutations, frameshifts, deletions, and insertions. Insertions (along with deletions and frameshifts) are more likely than point mutations to disrupt gene functions totally, and hence insertions may have more

severe deleterious effects, on average, than a random set of all mutations. Second, the approach used by Kibota and Lynch required that cells be able to form a visible colony within one day under the stringent growth conditions of their mutation-accumulation experiment; consequently, mutations with severe deleterious effects must have been excluded from their study. By contrast, we sought to include all mutations in our study by screening for mutations under permissive growth conditions before subsequently measuring their fitness under much more stringent conditions. If we exclude the handful of mutations with severe deleterious effects (i.e., the uniform component of the compound distribution), then the average selection coefficient is only 0.016. Thus, these two radically different approaches yield remarkably similar estimates for the average selection coefficient of *mildly* deleterious mutations.

Future directions

In the present study, we examined the fitness effects that were caused by single mutations. We are extending this research program to investigate the fitness effects produced by multiple mutations (Elena & Lenski, 1997). The same mini-Tn*10* derivatives employed in the present study are ideally suited to this approach because they encode three different markers for screening, which allow us to construct sets of genotypes that carry one, two, or three mutations. We can then use these genotypes to determine whether successive deleterious mutations typically have multiplicative effects on fitness or, alternatively, whether they tend to interact synergistically such that two mutations are more deleterious than expected from their separate effects. A systematic tendency for deleterious mutations to interact synergistically is a requirement of the mutational deterministic hypothesis (Kondrashov, 1993), which offers a relatively simple explanation for the adaptive significance of sexuality.

We are also using genotypes that have these mini-Tn*10* insertion mutations to examine evolutionary reversibility and historicity (F.B.-G. Moore, D.E. Rozen, S.F. Elena & R.E. Lenski, unpublished). If populations of these genotypes are propagated for hundreds of generations, will they achieve higher fitness by eliminating these insertions? Will they be saddled with these harmful effects indefinitely? Or will they achieve higher fitness via compensatory mutations that specifically relieve the harmful effects of the insertion mutations? And if compensatory mutations are commonplace, do *they* tend to have harmful side-effects? That is, if the original insertion mutations were removed, then would the compensatory mutations themselves be deleterious? The answers to these questions may shed some light on the structure of adaptive landscapes and the mechanisms of speciation (Wright, 1932, 1982, 1988).

Acknowledgements

We thank L. Forney, V. de Lorenzo, O. Martínez de Ilarduya and M. Pérez-Amador for valuable advice on experimental methods. We also thank P. Gerrish, J. Mongold, P. Moore, D. Rozen, P. Sniegowski, A. de Visser, and an anonymous reviewer for helpful comments on this manuscript. This research was supported by a fellowship from the Spanish Ministerio de Educación y Ciencia to S.F.E. and by a National Science Foundation grant to R.E.L.

References

Barton, N.H. & M. Turelli, 1987. Adaptive landscapes, genetic distance and the evolution of quantitative characters. Genetical Res. 49: 157–173.

Barton, N.H. & M. Turelli, 1989. Evolutionary quantitative genetics: how little do we know? Annu. Rev. Genet. 23: 337–370.

Butcher, D., 1995. Muller's ratchet, epistasis and mutation effects. Genetics 141: 431–437.

Charlesworth, D., M.T. Morgan & B. Charlesworth, 1993. Mutation accumulation in finite outbreeding and inbreeding populations. Genetical Res. 61: 39–56.

Chao, L., C. Vargas, B.B. Spear & E.C. Cox, 1983. Transposable elements as mutator genes in evolution. Nature 303: 633–635.

Clark, A.G., L. Wang & T. Hulleberg, 1995a. Spontaneous mutation rate of modifiers of metabolism in Drosophila. Genetics 139: 767–779.

Clark, A.G., L. Wang & T. Hulleberg, 1995b. *P*-element induced variation in metabolic regulation in Drosophila. Genetics 139: 337–348.

Crow, J.F. & M. Kimura, 1970. An Introduction to Population Genetics Theory. New York: Harper & Row.

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

Dykhuizen, D.E. & A.M. Dean, 1990. Enzyme activity and fitness: evolution in solution. Trends Ecol. Evol. 5: 257–262.

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

Elena, S.F. & R.E. Lenski, 1997. Test of synergistic interactions among deleterious mutations in bacteria. Nature 390: 395–398.

Gerrish, P.J. & R.E. Lenski, 1998. The fate of competing beneficial mutations in an asexual population. Genetica 102/103: 127–144.

Gregory, W.C., 1965. Mutation frequency, magnitude of change and the probability of improvement in adaptation. Radiation Botany 5 (suppl.): 429–441.

Haldane, J.B.S., 1927. A mathematical theory of natural and artificial selection. V. Selection and mutation. Proc. Camb. Phil. Soc. 23: 838–844.

Haldane, J.B.S., 1937. The effect of variation on fitness. Amer. Nat. 71: 337–349.

Hill, W.G. & J. Rasbash, 1986. Models of long-term artificial selection in finite population with recurrent mutation. Genetical Res. 48: 125–131.

Houle, D., D.K. Hoffmaster, S. Assimacopoulos & B. Charlesworth, 1992. The genomic mutation rate for fitness in *Drosophila*. Nature 359: 58–60.

Houle, D., B. Morikawa & M. Lynch, 1996. Comparing mutational variabilities. Genetics 143: 1467–1483.

Keightley, P.D., 1994. The distribution of mutation effects on viability in *Drosophila melanogaster*. Genetics 138: 1315–1322.

Keightley, P.D., 1996. Nature of deleterious mutation load in *Drosophila*. Genetics 144: 1993–1999.

Keightley, P.D. & W.G. Hill, 1988. Quantitative genetic variability maintained by mutation-stabilizing selection balance in finite populations. Genetical Res. 52: 33–43.

Kibota, T.T. & M. Lynch, 1996. Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. Nature 381: 694–696.

Kimura, M., 1979. Model of effectively neutral mutations in which selective constraint is incorporated. Proc. Natl. Acad. Sci., USA 76: 3440–3444.

Kimura, M., 1983. The Neutral Theory of Molecular Evolution. Cambridge: Cambridge University Press.

Kleckner, N., J. Bender & S. Gottesman, 1991. Uses of transposons with emphasis on Tn*10*. Meth. Enzymol. 204: 139–180.

Kondrashov, A.S., 1993. Classification of hypotheses on the advantage of amphimixis. J. Heredity 84: 372–387.

Kondrashov, A.S., 1994. Muller's ratchet under epistatic selection. Genetics 136: 1469–1473.

Lande, R., 1975. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genetical Res. 26: 221–235.

Lande, R., 1983. The response to selection on major and minor mutations affecting a metrical trait. Heredity 50: 47–65.

Lande, R., 1994. Risk of population extinction from fixation of new deleterious mutations. Evolution 48: 1460–1469.

Lenski, R.E., 1988. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. Variation in competitive fitness among mutants resistant to virus T4. Evolution 42: 425–432.

Lenski, R.E., 1992. Experimental evolution, pp. 125–140 in Encyclopedia of Microbiology, Vol. 2, edited by J. Lederberg. San Diego: Academic Press.

Lenski, R.E., M.R. Rose, S. C. Simpson & S.C. Tadler, 1991. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2000 generations. Amer. Nat. 138: 1315–1341.

Lenski, R.E. & M. Travisano, 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc. Natl. Acad. Sci. USA 91: 6808–6814.

Levin, B.R. & R.E. Lenski, 1983. Coevolution in bacteria and their viruses and plasmids, pp. 99–127 in Coevolution, edited by D. J. Futuyma & M. Slatkin. Sunderland, MA: Sinauer Associates.

Levin, B.R., F.M. Stewart & L. Chao, 1977. Resource-limited growth, competition, and predation: a model and experimental studies with bacteria and bacteriophage. Amer. Nat. 111: 3–24.

López, M.A. & C. López-Fanjul, 1993. Spontaneous mutation for a quantitative trait in *Drosophila melanogaster*. II. Distribution of mutant effects on the trait and fitness. Genetical Res. 61: 117–126.

Lyman, R.F., F. Lawrence, S.V. Nuzhdin & T.F.C. Mackay, 1996. Effects of single *P*-element insertions on bristle number and viability in *Drosophila melanogaster*. Genetics 143: 277–292.

Lynch, M., J. Conery & R. Bürger, 1995. Mutation accumulation and the extinction of small populations. Amer. Nat. 146: 489–518.

Lynch, M. & W. Gabriel, 1990. Mutation load and the survival of small populations. Evolution 44: 1725–1737.

Mackay, T.F.C., R.F. Lyman & M.S. Jackson, 1992. Effects of *P* element insertion on quantitative traits in *Drosophila melanogaster*. Genetics 130: 315–332.

Mukai, T., 1964. The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. Genetics 50: 1–19.

Mukai, T., S.I. Chigusa, L.E. Mettler & J.F. Crow, 1972. Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. Genetics 72: 333–335.

Muller, H.J., 1950. Our load of mutations. Amer. J. Human Genet. 2: 111–176.

Muller, H.J., 1964. The relation of recombination to mutational advance. Mutat. Res. 1: 2–9.

Nguyen, T.N.M., Q.G. Phan, L.P. Duong, K.P. Bertrand & R.E. Lenski, 1989. Effects of carriage and expression of the Tn*10* tetracycline resistance operon on the fitness of *Escherichia coli* K12. Mol. Biol. Evol. 6: 213–225.

Ohta, T., 1977. Extensions to the neutral mutation random drift hypothesis, pp. 148–167 in Molecular Evolution and Polymorphism, edited by M. Kimura. Mishima, Japan: National Institute of Genetics.

Press, W.H., B.P. Flannery, S.A. Teukolsky & W.T. Vetterling, 1990. Numerical Recipes in Pascal. Cambridge: Cambridge University Press.

Sniegowski, P.D. & R.E. Lenski, 1995. Mutation and adaptation: the directed mutation controversy in evolutionary perspective. Ann. Rev. Ecol. Syst. 26: 553–578.

Sniegowski, P.D., P.J. Gerrish & R.E. Lenski, 1997. Evolution of high mutation rates in experimental populations of *E. coli*. Nature 387: 703–705.

Sokal, R.R. & F.J. Rohlf, 1981. Biometry, $2^{n\,d}$ edition. New York: W. H. Freeman.

Turelli, M., 1984. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theor. Pop. Biol. 25: 138–193.

Wright, S., 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. Sixth. Intl. Congr. Genet. 1: 356–366.

Wright, S., 1982. Character change, speciation, and the higher taxa. Evolution 36: 427–443.

Wright, S., 1988. Surfaces of selective value revisited. Amer. Nat. 131: 115–123.