

# *Escherichia coli* rpoB Mutants Have Increased Evolvability in Proportion to Their Fitness Defects

Jeffrey E. Barrick,\* Mark R. Kauth Christopher C. Strelhoff, and Richard E. Lenski

Department of Microbiology and Molecular Genetics, Michigan State University

\*Corresponding author: E-mail: jbarrick@msu.edu.

Associate editor: Helen Pionthkivska

## Abstract

Evolvability is the capacity of an organism or population for generating descendants with increased fitness. Simulations and comparative studies have shown that evolvability can vary among individuals and identified characteristics of genetic architectures that can promote evolvability. However, little is known about how the evolvability of biological organisms typically varies along a lineage at each mutational step in its history. Evolvability might increase upon sustaining a deleterious mutation because there are many compensatory paths in the fitness landscape to reascend the same fitness peak or because shifts to new peaks become possible. We use genetic marker divergence trajectories to parameterize and compare the evolvability—defined as the fitness increase realized by an evolving population initiated from a test genotype—of a series of *Escherichia coli* mutants on multiple timescales. Each mutant differs from a common progenitor strain by a mutation in the *rpoB* gene, which encodes the  $\beta$  subunit of RNA polymerase. Strains with larger fitness defects are proportionally more evolvable in terms of both the beneficial mutations accessible in their immediate mutational neighborhoods and integrated over evolutionary paths that traverse multiple beneficial mutations. Our results establish quantitative expectations for how a mutation with a given deleterious fitness effect should influence evolvability, and they will thus inform future studies of how deleterious, neutral, and beneficial mutations targeting other cellular processes impact the evolutionary potential of microorganisms.

**Key words:** evolvability, compensatory adaptation, fitness landscape, *Escherichia coli*, experimental evolution, neutral marker divergence.

## Introduction

Mutations can affect not only the current fitness of an organism but also the ability of its descendants to evolve and adapt by natural selection. This capacity for improvement is known as “evolvability.” It has been proposed that differences in evolvability have favored the success of certain biological phyla and that design principles such as modularity and robustness contribute to increased evolvability (Wagner and Altenberg 1996; Kirschner and Gerhart 1998). In addition, studies with computational models have demonstrated that genetic architectures with greater evolvability can arise and will be favored under certain conditions (Earl and Deem 2004; Draghi and Wagner 2008). However, the question of whether evolvable genetic architectures in biological organisms result from clade-level selection or are a by-product of mutations with immediate benefits to individuals remains controversial (Leroi 2000; Sniegowski and Murphy 2006).

Evolvability, defined here as the fitness increase realized by an evolving population initiated from a test genotype, is a complex trait that is difficult to measure. It depends on how mutational processes generate genetic variation, how developmental and regulatory processes render these changes into potentially adaptive phenotypic variation, the population dynamics of competition between contending beneficial mutations, and whether the timescale of interest spans a single

mutational step or many steps. These factors can be rigorously controlled and tested with replication in evolution experiments with microorganisms (Colegrave and Collins 2008). Such studies have shown that RNA virus genotypes with similar fitness may differ in their evolvability (Burch and Chao 2000; McBride et al. 2008) and that bacteria with increased mutation rates may be more evolvable under certain circumstances (Sniegowski et al. 1997; de Visser et al. 1999).

Little is known about how different kinds of mutations impact microbial evolvability and whether there are any general principles for predicting these effects. To begin to address these questions about the fitness landscapes of biological organisms, we measured the evolvability of a series of *Escherichia coli* strains with different *rpoB* mutations on two timescales. On short timescales, when the first beneficial mutations sweep through evolving populations, and over longer periods, when multiple beneficial mutations accrue, we find a strikingly constant relationship between the fitness defects caused by these mutations and the degree to which they increase a strain's evolvability.

## Materials and Methods

### Test Strains

Eight rifampicin-resistant (Rif<sup>R</sup>) mutants of the reference strain REL606 (Lenski et al. 1991; Jeong et al. 2009) were selected in a single step on Luria Bertani (LB) plates

supplemented with 100  $\mu\text{g/ml}$  rifampicin. We cannot rule out the possibility that unknown second-site mutations occurred during the isolation of these strains, but there is generally thought to be an extremely small chance of recovering double mutations with this procedure (Kassen and Bataillon 2006; MacLean and Buckling 2009), and the fitness defects we found are similar to those reported in other studies (Reynolds 2000). Variants of each test strain differing in a genetic marker for L-arabinose (Ara) utilization were constructed for evolution and competition experiments. The reference strain has an inactivating point mutation in the *araA* gene that gives it an Ara<sup>−</sup> phenotype (Lenski et al. 1991). Ara<sup>+</sup> revertants of each test strain were isolated by selective plating or gene gorging (Herring et al. 2003). Ara<sup>−</sup> and Ara<sup>+</sup> strains can be distinguished as red and white colonies, respectively, on tetrazolium arabinose (TA) indicator agar (Lenski et al. 1991).

### Evolution Experiment

We propagated 12 replicate populations of each test strain (the reference strain and eight Rif<sup>R</sup> mutants) founded from equal mixtures of its Ara<sup>−</sup> and Ara<sup>+</sup> marked variants for 80 days as 3-ml liquid cultures in Davis minimal media (Lenski et al. 1991) supplemented with 100  $\mu\text{g/l}$  glucose (DM100). To ensure that beneficial mutations in each evolutionary replicate were independently derived, we started each replicate from two separate cultures, one of the Ara<sup>−</sup> variant and one of the Ara<sup>+</sup> variant, each inoculated with only 30–50 cells. After 48 h of growth to saturating density, the generation zero population was established by diluting each paired Ara<sup>−</sup> and Ara<sup>+</sup> culture 1:512 into the same 3-ml test tube of fresh media. Thereafter, every 24 h, we transferred 11.7  $\mu\text{l}$  of each culture into 3 ml fresh growth media for a 1:256 dilution, allowing eight generations of growth to a final population size of  $\sim 5.7 \pm 2.3 \times 10^8$  cells (95% confidence interval). Cultures were incubated at 37 °C in 18  $\times$  150-mm test tubes with orbital shaking over a diameter of 1.9 or 2.5 cm at 160 rpm. Every 2 days (16 generations), we plated a dilution from each of the 108 populations on TA agar and counted 200–600 colonies to determine the ratio of the two marked strains. When the marker ratio in a population diverged by more than 100-fold in either direction for two consecutive measurements, we ceased plating and counting that replicate until the end of the experiment.

The final Ara<sup>−</sup>/Ara<sup>+</sup> ratio was measured again on day 80 for all populations, at which time none were dominated by a different color than when plating had been suspended. We isolated a clone with the majority marker state from each final population and sequenced the portion of its *rpoB* gene containing the original Rif<sup>R</sup> mutation to test for contamination. One Rif<sup>R</sup>-8 replicate had the wrong *rpoB* mutation and was excluded from all analyses. We did not observe any reversions of the original *rpoB* mutations responsible for resistance, and all evolved Rif<sup>R</sup> clones remained resistant to 100  $\mu\text{g/ml}$  rifampicin when streaked on LB plates. We found secondary mutations that may contribute to adaptation in the *rpoB* region we sequenced in 5 of the 12 final Rif-7 clones, but none in the other evolved clones.

The marker trajectories for the Rif<sup>R</sup>-6 and reference strains trended consistently toward one color early in the evolution experiment across all replicates (see supplementary figs. S1 and S2, Supplementary Material online), indicating that their Ara<sup>−</sup> and Ara<sup>+</sup> variants did not have precisely the same fitness. We estimate that the Ara<sup>−</sup> Rif<sup>R</sup>-6 strain had a fitness advantage of  $1.28\% \pm 0.23\%$  over its paired Ara<sup>+</sup> revertant during the first 64 generations of the experiment and that the Ara<sup>+</sup> variant of the reference strain had a  $0.55\% \pm 0.20\%$  advantage over the Ara<sup>−</sup> version during the first 80 generations ( $n = 12$ , standard error of the mean). These differences may be due to unknown second-site mutations sustained while constructing Ara<sup>+</sup> variants or during their propagation prior to the evolution experiment. In any case, their fitness effects are negligible compared with the selection coefficients of the beneficial mutations studied here, and we also corrected for these trending baselines in the marker trajectory analysis.

### Marker Divergence Analysis

The procedure for inferring effective evolutionary parameters for the first beneficial mutations to fix in a test strain from replicate marker ratio trajectories, following Hegreness et al. (2006), has three main steps: 1) Simulate families of marker trajectories under the same transfer regime as the evolution experiment at many combinations of  $\mu$  and  $s$  parameters that describe the mutation rate and selection coefficient, respectively, of a single category of beneficial mutations. 2) Fit the shape of the initial divergence of each experimental and simulated marker trajectory to an empirical equation with  $\alpha$  and  $\tau$  parameters that represent the rate of divergence and the waiting time until divergence, respectively. 3) Find the values of the effective evolutionary parameters  $\mu$  and  $s$  at which the 2D distribution of ( $\alpha$ ,  $\tau$ ) pairs fit to the simulated trajectories best reproduces the ( $\alpha$ ,  $\tau$ ) pairs fit to the experimental trajectories. Details for how we carried out each of these steps in the analysis are described below.

**Stochastic Population Genetic Simulations.** Hegreness et al. (2006) modeled the evolutionary dynamics of serial transfer experiments with a combination of deterministic and stochastic methods. In this work, we employ a continuous time Markov process for all growth and mutation dynamics (Renshaw 1993) and model dilution as sampling without replacement, making all aspects of these simulations stochastic. This approach means that we do not employ any of the assumptions needed to efficiently combine deterministic and stochastic methods. In particular, the first beneficial mutations to fix in some Rif strain backgrounds have large selection coefficients that would invalidate approximations that were used previously to calculate the probability that a newly generated beneficial mutation establishes in a population of fluctuating size in the deterministic model (Wahl and Gerrish 2001).

Our simulation tracks the number of cells with a given neutral marker background  $b$  and number of beneficial mutations  $m$  at time  $t$ :  $N_{b,m}(t)$ . The marker background

can be red or white:  $b \in \{r, w\}$ ; and the number of mutations can range from zero to three:  $m \in \{0, 1, 2, 3\}$ . The base birth rate is given by  $r = \log_e 2$ , the doubling time of cells without mutations. Each mutation gives an additive selective benefit of constant size  $s$ . As a result, cells with  $m$  mutations will have a birth rate of  $r(1 + ms)$ . The rate of beneficial mutations per generation is given by  $\mu$ . Thus, the rates of birth and mutation are

$$N_{b,m} \rightarrow N_{b,m} + 1 : r(1 + ms)(1 - \mu)N_{b,m}(t), \quad (1)$$

$$N_{b,m+1} \rightarrow N_{b,m+1} + 1 : r(1 + ms)\mu N_{b,m}(t), \quad (2)$$

for allowed values of  $b$  and  $m$ . However, note that equation (2) is only valid for  $m \leq 2$  because we allow a maximum of three mutations.

Given this setup, the simulations proceed very much like the actual experiments. The parameters that must be set are the selective benefit of mutations  $s$  and the rate of beneficial mutations  $\mu$ . Once these are set, all simulations start with an initial growth phase. In this step, 100 cells with each marker background are grown separately to a final population size of  $N_f = 7.6 \times 10^8$ . Both populations are diluted by a factor of 1:512 and then mixed to create an initial population of size  $N_s = 2,968,750$ . From this point forward, the combined red and white population is grown from  $N_s$  to  $N_f$  and diluted by a factor of 1:256, constituting one transfer. Simulations continued until 1) 125 transfers were completed (1,000 generations) or 2) the absolute value of  $\log_{10}(N_r/N_w)$  was greater than two, indicating a 100-fold numerical advantage for one of the marker states.

The details of the computational implementation are as follows: Birth and mutation were simulated using a combination of the optimized tau-leaping method and basic stochastic simulation algorithm (SSA) as described in Cao et al. (2006). Tau leaping provides a great speed increase for these simulations and was used until the population size was near  $N_f$ . At this point, simulation switched to the exact SSA method to avoid exceeding  $N_f$ . Dilution was done by using the sample without replacement algorithm of Bebbington (1975). Finally, great care was taken with random number generation. We employed the Mersenne Twister algorithm (Matsumoto and Nishimura 1998) to generate uniform deviates and the algorithm of Ahrens and Dieter (1982) for Poisson deviates used in the tau-leaping algorithm.

Note that these population genetic simulations capture the important effects of clonal interference. Multiple mutational events generating each category of mutant can occur in a single simulation, and multiple subpopulations with different beneficial mutations typically compete before one marker state is fixed. The main simplification of this approach, justified in Hegreness et al. (2006), is that every beneficial mutation, even subsequent mutations in backgrounds that already have one or two beneficial mutations, has the same selection coefficient and rate of appearance. One generally expects secondary and tertiary mutations to be less beneficial than the first mutations to

sweep because there are fewer ways to improve the closer one comes to a fitness peak. However, this simplifying assumption does not invalidate the predictions of this approach because secondary and tertiary mutations do not appreciably affect the initial divergence of marker trajectories under our conditions. They mostly affect only the later dynamics in the marker trajectories, such as trend reversals, that are not part of the quantitative model used to fit the empirical data.

### Fitting the Initial Divergence of Marker Trajectories.

We fit the initial portion of each experimental and simulated marker trajectory to a simple exponential model with empirical parameters  $\alpha$  and  $\tau$  that represent the rate of divergence and the waiting time until divergence, respectively. Our procedure is an extension of that used by Hegreness et al. (2006). We include additional variables in the model to correct for experimental complications. Specifically, the R statistics package (version 2.6.1) (R Development Core Team 2009) was used to calculate nonlinear least squares fits of the initial divergence of marker trajectories to the following equation when the Ara<sup>-</sup> (red) variant won:

$$\log_e(R_{r/w}(t)) = \log_e(R_{r/w}(0)) + \Delta w_{r/w}t + \log_e(1 + F_r e^{\alpha(t-\tau)}), \quad (3)$$

$$F_r = R_{r/w}(0)/(1 + R_{r/w}(0)), \quad (4)$$

where  $R_{r/w}(t)$  is the ratio of red (Ara<sup>-</sup>) to white (Ara<sup>+</sup>) colonies at transfer  $t$ ,  $\Delta w_{r/w}$  is the initial fitness advantage of the red variant of the test strain over the white variant expressed as an additive selection coefficient per transfer, and  $R_{r/w}(0)$  is the initial ratio of red to white colonies at time 0. When a marker trajectory diverged in the opposite direction, such that the Ara<sup>+</sup> (white) variant prevailed, the same equation was used with  $r$  and  $w$  subscripts reversed.

For fitting simulated marker curves, we omitted the  $\Delta w_{r/w}$  parameter and set  $R_{r/w}(0) = 1$ . For experimental trajectories,  $R_{r/w}(0)$  and  $\Delta w_{r/w}$  were determined by fitting a user-specified number of initial baseline points for each marker trajectory to the equation:  $\log_e(R_{r/w}(t)) = \log_e(R_{r/w}(0)) + \Delta w_{r/w}t$ . Values of  $R_{r/w}(0)$  varied slightly due to imperfect mixing of the strains, which led to deviations from the intended 1:1 initial ratio in some replicates. A nonzero  $\Delta w_{r/w}$  value indicates that a pair of marked test strain variants had unequal fitness. Except for the reference and Rif<sup>R</sup>-6 strain pairs, where the replicate marker trajectories consistently trended toward one marker state (see [supplementary figs. S1 and S2](#), [Supplementary Material](#) online), the fit value of  $\Delta w_{r/w}$  was generally not significantly different from zero.

The procedure for finding the initial divergence parameters proceeds by testing the first points of each marker trajectory until there is evidence that adding the next point results in a data set that does not fit the model. The criteria for adding further points were that 1) the residual standard error of the fit was  $\leq 0.15$  and 2) the hypothesis that the residuals are normally distributed was not rejected by the

Lilliefors test at a 0.05 significance level. For some experimental trajectories, the standard error condition was never met due to excessive noise in the marker ratio measurements. In these cases, we used the set of initial points where the fit to the empirical model passed the Lilliefors test with the minimum standard error.

**Inferring Effective Evolutionary Parameters.** In order to determine the values of the effective evolutionary parameters  $\mu$  and  $s$  where simulations reproduce experimental marker trajectories, Hegreness et al. (2006) combined two 1D Kolmogorov–Smirnov tests that separately compared the  $\alpha$  and  $\tau$  distributions fit from the simulations and experiments. This procedure assumes that the distributions of  $\alpha$  and  $\tau$  are independent. Although this appears to be true for small values of  $s$ , it is not a good assumption for some of the larger  $s$  values encountered here (see, e.g., fig. 1c). Therefore, we used instead a brute force implementation of the Fasano and Franceschini (1987) algorithm for computing a 2D Kolmogorov–Smirnov test. Values of the  $D$  statistic for this test were converted to  $P$  values for rejecting the null hypothesis that the distributions are the same by using the approximation in Press et al. (2007).

Specifically, we simulated a total of 200 marker divergence trajectories for each of 806 combinations of  $\mu$  and  $s$ :  $-8.5 \leq \log_{10}(\mu) \leq -6.0$  in  $\log_{10}$  increments of 0.1 and  $0.05 \leq s \leq 0.35$  in increments of 0.01. Because of the relatively limited number of marker trajectories that could be simulated at each  $(\mu, s)$  combination due to computational time constraints, we performed an ad hoc smoothing procedure. The experimental distribution of  $(\alpha, \tau)$  parameter pairs for a given test strain was compared with a composite distribution that combined the  $(\alpha, \tau)$  values from simulations at the  $(\mu, s)$  combination in question with those from the eight surrounding families of simulations with adjacent values of  $\mu$  and  $s$ . That is, 1,800 pseudo-replicated  $(\alpha, \tau)$  pairs from neighboring simulations were used as the theoretical distribution for a given  $(\mu, s)$  combination. This procedure was used to determine 95% confidence intervals for  $\mu$  and  $s$  where the hypothesis that the simulated and experimental data were the same could not be rejected at a 0.05 significance level. Combinations of  $\mu$  and  $s$  that gave the best agreement were defined as those that maximized the 2D Kolmogorov–Smirnov test  $P$  value.

As in Hegreness et al. (2006), we assume that single mutants are almost always responsible for the initial divergence of marker trajectories. It is possible for the beneficial mutation rate to be so high that, by the time there is appreciable divergence, multiple small mutations have usually occurred. This possibility naturally gives rise to a second region of agreement in the  $(\mu, s)$  plane that corresponds to a lower selection coefficient and higher mutation rate. The edge of this area shows up in figure 1d for Rif<sup>R</sup>-3 but is outside the plotted area for all other test strains. We can be confident that single mutants are responsible for divergence, and thus discount this possibility in our experiments, because we know that the overall mutation rate is very low for our strain (Barrick et al. 2009). However, this discontinuous confidence

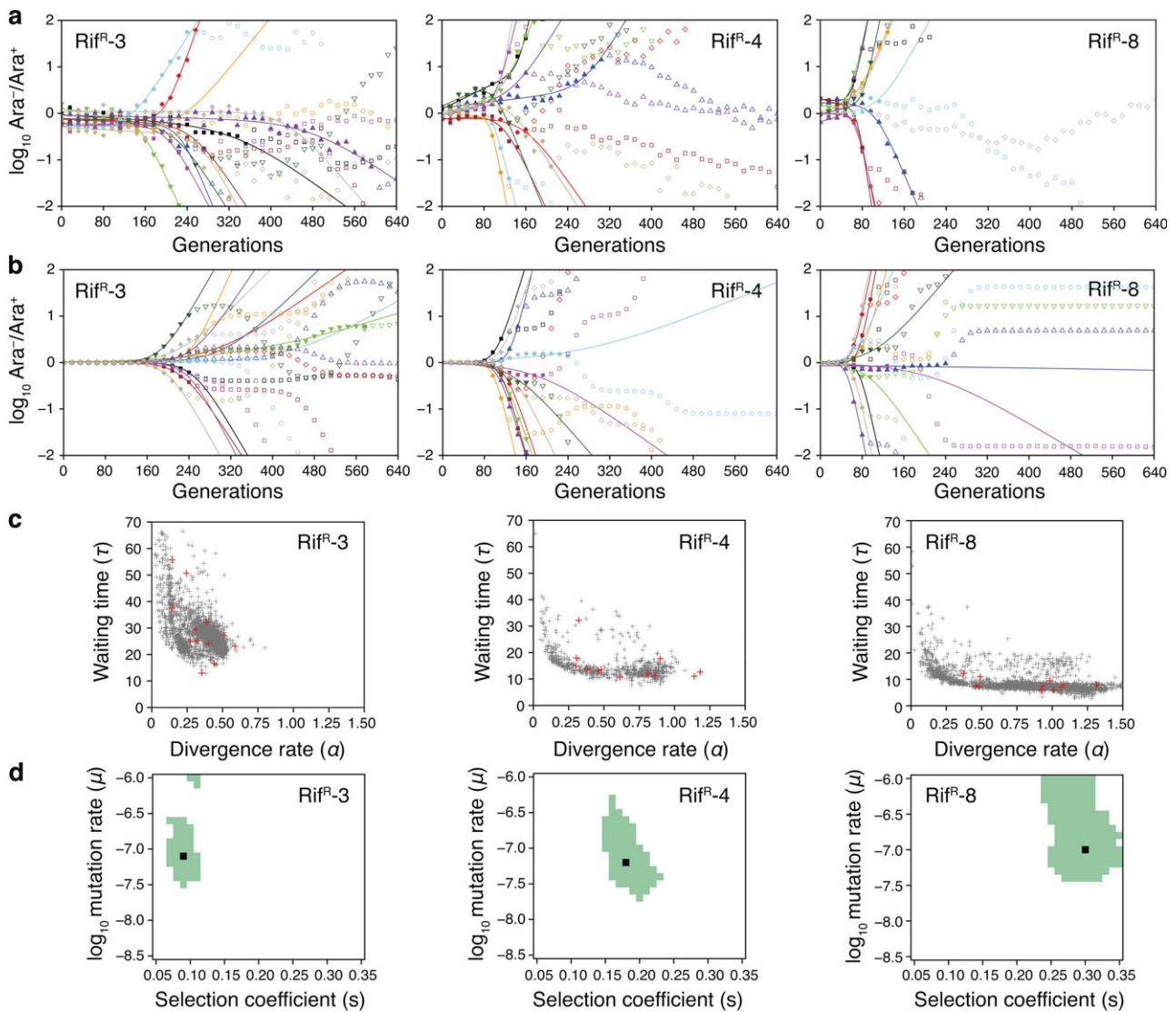
interval may apply in other systems. The opposite extreme is also possible. In some evolving systems, it may be necessary for substantial drift on a neutral network to occur before any beneficial mutations can be discovered, or small beneficial steps may make beneficial mutations of much larger effect possible. In these cases, the initial divergence of marker trajectories will also not be due to a single mutant. We believe that these cases would lead to predictions of very low effective beneficial mutation rates (far below the reciprocal of the effective population size and unlike what we found in our experiment).

### Fitness Measurements

Initial fitness defects and evolved improvements were determined using competition experiments (Lenski et al. 1991). Each pair of Ara<sup>-</sup> and Ara<sup>+</sup> strains being tested was separately revived from stocks stored at  $-80^{\circ}\text{C}$  in 15% glycerol or 15% dimethyl sulfoxide, under the same conditions as the evolution experiment, for one 24-h growth cycle. After a 1:256 dilution into separate test tubes for each replicate measurement and an additional 24-h growth cycle for acclimation to the experimental conditions, we diluted each Ara<sup>-</sup> and Ara<sup>+</sup> pair 1:512 into the same 3-ml test tube of DM100 and immediately spread an appropriate dilution in saline solution onto TA agar to determine the initial representation of each strain. After a 24-h growth cycle in coculture, an appropriate dilution was again plated on TA agar to determine the final representation of each strain. Relative fitness was calculated as the ratio of the realized Malthusian parameters for each strain (i.e., the ratio of the logarithm of the final over initial cell numbers for each strain) (Lenski et al. 1991).

### Test of Fitness Transitivity

Nontransitive fitness interactions (Paquin and Adams 1983) between strains bearing the deleterious Rif<sup>R</sup> mutations and the beneficial mutations that arise in the evolution experiment could lead to erroneous relative fitness values after normalizing to the reference strain's fitness. Therefore, we checked this assumption by making three sets of fitness measurements for each test strain: the relative fitness of the marked test strain variants versus the reference strain (six measurements per marker state, total  $n = 12$ ), the average relative fitness of each test strain versus representative evolved clones isolated from each of its 12 experimental populations at day 80 (one measurement per population, total  $n = 12$ ), and the average relative fitness of the 12 representative clones versus the reciprocally marked reference strain (one measurement per population, total  $n = 12$ ). For each of the eight test strains, the average fitness of its evolved clones measured directly relative to the reference strain agreed with that predicted after combining measurements of its evolved clones versus the test strain and the test strain versus the reference strain (two-tailed Welch's paired  $t$ -test, all  $P > 0.05$ ). Thus, there is no evidence for nontransitive fitness interactions that would complicate our analysis.



**FIG. 1.** Single-step evolvability inferred from marker ratio divergence trajectories. The  $Rif^R$ -3,  $Rif^R$ -4, and  $Rif^R$ -8 test genotypes illustrate the procedure for estimating the effective evolutionary parameters for the first beneficial mutations that sweep to high frequency in a test strain. (a) Experimental marker trajectories. The ratio of two marked variants ( $Ara^-/Ara^+$ ) of each test strain was monitored during a 640-generation evolution experiment. Each set of colored symbols represents an independently evolving replicate population (12 each for  $Rif^R$ -3 and  $Rif^R$ -4 and 11 for  $Rif^R$ -8). The marker ratio diverges from unity as beneficial mutations that occur in the genetic background of one marker state rise in frequency during the evolution experiment. Curves are fits to an empirical equation with parameters describing the waiting time ( $\tau$ ) and steepness ( $\alpha$ ) of the initial divergence. Fits include only the initial points of each marker trajectory (closed symbols) until goodness-of-fit tests fail (open symbols). In large populations such as these, the shape of initial divergence is often not attributable to a single beneficial mutation. Rather, it typically represents the superimposed effects of clonal interference between multiple competing beneficial mutations, some linked to each of the two marker states. In particular, certain marker curves that show delayed divergence may indicate that almost equally beneficial mutations arose nearly simultaneously in each of the two marker backgrounds. (b) Simulated marker trajectories. To infer a characteristic effective per-generation mutation rate ( $\mu$ ) and selective advantage ( $s$ ) for the first successful beneficial mutations in a given strain background, we used stochastic population genetic simulations that include clonal interference to generate 200 theoretical marker divergence curves for each of 806 different ( $\mu$ ,  $s$ ) combinations. The initial divergence of these trajectories was fit to the empirical equation to obtain a distribution of ( $\alpha$ ,  $\tau$ ) parameter pairs describing this family of curves. Twelve simulated marker trajectories, generated using  $\mu$  and  $s$  values with the best agreement to the experimental data, are shown with initial divergence curves fit as in (a). (c) Empirical parameters. The distributions of  $\alpha$  and  $\tau$  empirical parameter pairs fit from experimental data (red crosses) and simulations (gray crosses) were compared using a 2D Kolmogorov–Smirnov test to reject those ( $\mu$ ,  $s$ ) combinations where the initial divergence statistics differ significantly ( $P < 0.05$ ) from the experimental data. The simulated distributions were generated using the  $\mu$  and  $s$  values that provide the best agreement to the experimental data.  $\tau$  values are expressed in units of transfers, and  $\alpha$  values are per transfer, where one transfer is equal to eight generations. (d) Single-step evolvability. The effective evolutionary parameters  $\mu$  and  $s$  measure the evolvability of each test strain in terms of the first beneficial mutations that sweep to fixation or near fixation under these conditions. Parameter combinations giving the best agreement between simulated and observed marker trajectories (black) and an estimated 95% confidence interval (green) are shown. Values of  $\mu$  are per cell generation, and values of  $s$  are additive selection coefficients normalized to the fitness of each  $Rif^R$  ancestor. The edge of a discontinuous confidence interval for  $Rif^R$ -3 appears on the upper border of its graph. This region corresponds to a very high beneficial mutation rate that is not relevant to our experiment, as discussed in the text. See [supplementary figures S1 and S2](#), Supplementary Material online, for graphs of the other six test strains, and see the Materials and Methods for full details of the analysis procedure.

**Table 1.** Test Strains of *Escherichia coli* with Rifampicin Resistance Mutations.

Isolate	Relative Fitness <sup>a</sup>	Mutated Gene <sup>b</sup>	Nucleotide Change <sup>c</sup>	Codon Change	Amino Acid Change <sup>c</sup>
Reference	1.004 ± 0.028	—	—	—	—
Rif <sup>R</sup> -1	0.984 ± 0.025	<i>rpoB</i> (II)	A1714C	ATC → CTC	I572L
Rif <sup>R</sup> -2	0.980 ± 0.036	<i>rpoB</i> (II)	C1721A	TCT → TAT	S574Y
Rif <sup>R</sup> -3	0.969 ± 0.028	<i>rpoB</i> (I)	G1546T	GAC → TAC	D516Y
Rif <sup>R</sup> -4	0.839 ± 0.018	<i>rpoB</i> (—)	A443T	CAG → CTG	Q148L
Rif <sup>R</sup> -5	0.806 ± 0.023	<i>rpoB</i> (I)	A1547G	GAC → GGC	D516G
Rif <sup>R</sup> -6	0.742 ± 0.018	<i>rpoB</i> (I)	C1535T	TCT → TTT	S512F
Rif <sup>R</sup> -7	0.728 ± 0.016	<i>rpoB</i> (I)	A1538C	CAG → CCG	Q513P
Rif <sup>R</sup> -8	0.698 ± 0.020	<i>rpoB</i> (I)	Δ1594–1605	—	Δ532–535

<sup>a</sup> Fitness relative to the reference strain (Ref) was measured using competition assays in the environment of the evolution experiment, which does not contain rifampicin. Errors are 95% confidence intervals estimated from replicate measurements ( $n = 12$ ).

<sup>b</sup> Mutations within common Rif resistance clusters (I and II) of the *rpoB* gene are indicated (Jin and Zhou 1996).

<sup>c</sup> Nucleotide and amino acid changes are relative to the coordinates of the *E. coli* *rpoB* gene.

### Linear Regression

Because there is experimental uncertainty in the test strain fitness defects ( $d$ ), we used Model II regression to test for relationships between  $d$  and the various evolvability parameters. The major axis method was used to fit regression lines against both the test strains' effective selection coefficients ( $s$ ) and their sustained fitness increases ( $\Sigma$ ). We estimated  $P$  values for how significantly the slopes were greater than zero and less than one by a method appropriate for small samples (Jolicoeur 1968). No significant correlation was found between  $d$  and the effective mutation rate ( $\mu$ ) or its base-ten logarithm by either major axis or ordinary least squares regression.

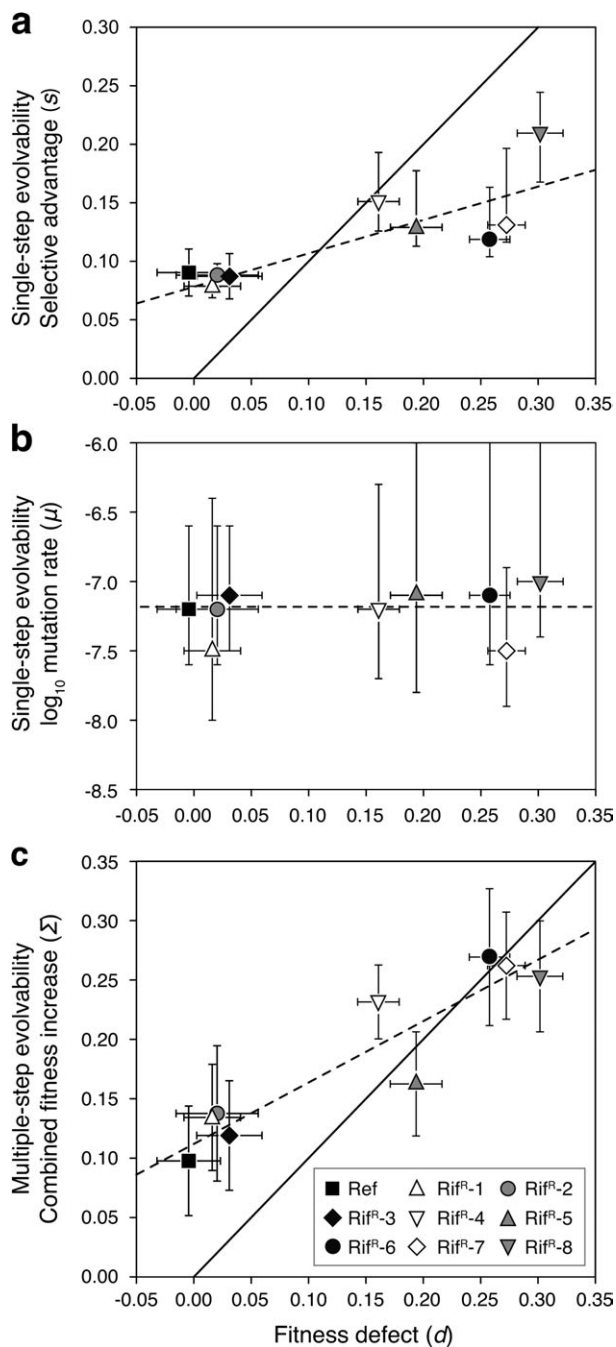
### Results and Discussion

To explore the effects of defined single-mutation steps on bacterial evolvability, we isolated Rif<sup>R</sup> variants of an *E. coli* B strain that has been the focus of other evolution studies (Lenski et al. 1991; Blount et al. 2008; Barrick et al. 2009). Most Rif<sup>R</sup> mutations in bacteria occur within the *rpoB* gene (Jin and Zhou 1996), which encodes the  $\beta$  subunit of RNA polymerase. Because Rif<sup>R</sup> mutants are at an advantage relative to their sensitive progenitors when the antibiotic rifampicin is present, they have been used previously to infer the distributions of fitness effects of beneficial mutations in these stressful environments (Kassen and Bataillon 2006; MacLean and Buckling 2009). Some Rif<sup>R</sup> mutations disrupt RNA polymerase function, perturbing global gene expression with pleiotropic and potentially maladaptive consequences. Thus, in the environment used in our experiments, which does not contain rifampicin, they represent neutral or deleterious mutations, such as those that might fix in an evolving population by genetic drift or hitchhiking with beneficial mutations. The eight Rif<sup>R</sup> isolates studied here have different mutations in *rpoB* and a range of fitness defects relative to the reference strain (table 1).

We performed evolution experiments in a marker divergence format (Chao and Cox 1983; Rozen et al. 2002; Hegreness et al. 2006) to quantify how each *rpoB* mutation affects evolvability. Twelve replicate populations were founded as equal mixtures of two variants of each test strain that were distinguishable by a neutral (or nearly neutral) genetic marker and propagated by serial transfer for 80 days

(640 generations) of evolution. Over time, new beneficial mutations will arise in the genetic background of one of the two marker states and cause the proportion of individuals with that marker state to increase as they displace less fit competitors of the opposite marker state and their ancestors (fig. 1a). Reproduction is strictly asexual in this system (Lenski et al. 1991), so mutations remain linked to the marker state in which they originate and there are no hybrid progeny that recombine beneficial mutations from competing lineages. In large asexual populations, such as the ones in this experiment, multiple subpopulations of variants carrying different beneficial mutations will typically reach relatively high frequencies in a population before one sweeps to fixation. Competition between these beneficial mutations, called clonal interference, slows the takeover of the population by the most fit subpopulation. As each beneficial mutation may occur in either of the two marker backgrounds, changes in the marker ratio over time represent a superposition of the effects of this competition, and therefore, these changes do not directly reflect the selective advantage of a single sweeping beneficial mutation relative to its progenitor.

Although there are many more mutations that confer small than large fitness advantages (Perfeito et al. 2007), those rare mutations with large benefits will typically prevail and be the first mutations to appreciably shift the marker ratio and sweep to fixation or near fixation in large populations (Gerrish and Lenski 1998; Patwa and Wahl 2008). In fact, a simplified model with only a single category of beneficial mutations with one effective mutation rate ( $\mu$ ) and selective advantage ( $s$ ) that also includes clonal interference between independently arising mutants can reasonably reproduce the evolutionary dynamics observed in these experiments (Hegreness et al. 2006). We compared the initial divergence in each set of marker trajectories with population genetic simulations to infer the underlying  $s$  and  $\mu$  values characteristic of the first beneficial mutations to sweep in populations founded by each test strain (fig. 1 and supplementary figs. S1 and S2, Supplementary Material online). These two effective parameters provide quantitative measures of short-term evolvability. They reflect the features of a genotype's local fitness landscape that determine its ability to adapt in a single step, given the environment and population structure.

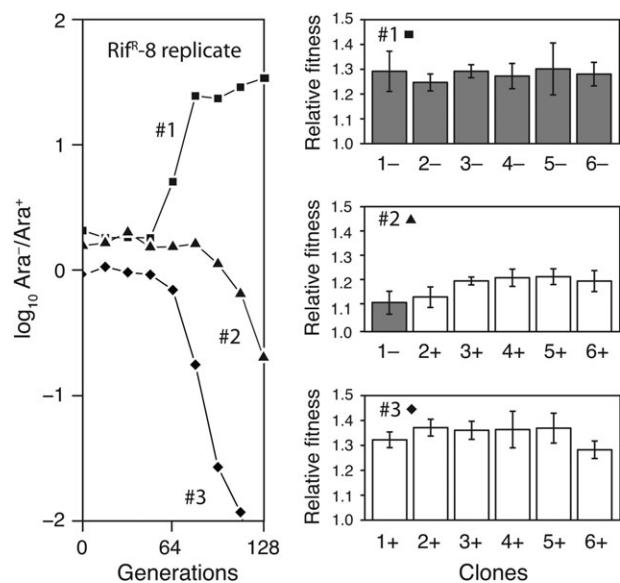


**Fig. 2.** Trends in single-step and multiple-mutation evolvability. The effects of deleterious Rif<sup>R</sup> mutations and evolved beneficial mutations are reported as additive selection coefficients normalized to the fitness of the reference *Escherichia coli* strain. Thus, the competitive fitness of an evolved Rif<sup>R</sup> strain relative to the reference strain is one minus its initial fitness defect ( $d$ ) plus the first successful mutation's selective advantage ( $s$ ) or plus the net fitness increase over multiple mutations ( $\Sigma$ ). (a) The effective selective advantage ( $s$ ) inferred for the first beneficial mutations in each test strain plotted against that strain's initial fitness defect ( $d$ ). Replicate competition assays ( $n = 12$ ) between each  $rpoB$  mutant and the ancestor strain were used to measure  $d$  (see Table 1). Values of  $s$  were inferred from a statistical analysis of marker divergence trajectories (see fig. 1). Error bars are 95% confidence intervals in both cases. The major axis regression line (dashed) and the line of equal fitness corresponding to the reference *E. coli* strain (solid) are shown. Values of  $s$  differ from those in figure 1 because they are

Beneficial mutations conferring larger selective advantages are accessible to test strains with lower initial fitness (fig. 2a). Strains with neutral or only slightly deleterious  $rpoB$  mutations evolve initial beneficial mutations with fitness effects similar to the reference strain, whereas the strains with the lowest initial fitness acquire beneficial mutations that confer at least twice this advantage. This overall trend is significant: Linear regression indicates a positive slope for the relationship ( $P = 0.003$ ) such that a strain's initial fitness defect is highly predictive of the magnitude of its first successful beneficial mutations ( $r^2 = 0.67$ ). However, the slope of this relationship is also much less than one ( $P = 0.0001$ ), indicating that the increase in evolvability for a strain with a deleterious  $rpoB$  mutation is not enough to compensate wholly for the defect. On average, the predicted difference in relative fitness between evolved Rif<sup>R</sup> and evolved reference strains, each with a single beneficial mutation, is 66% of the difference between their ancestors. The fitness increases measured directly by competing clones isolated from three Rif<sup>R</sup>-8 populations at generation 128 (just as the first beneficial mutations near fixation) against their Rif<sup>R</sup> ancestor corroborate the selection coefficients inferred from the marker divergence analysis (fig. 3).

In contrast, the effective rate ( $\mu$ ) at which the first successful beneficial mutations occur in a test strain does not depend on that strain's fitness defect (fig. 2b). The best fit slope for this relationship is not significantly different from zero (see Materials and Methods). The average  $\mu$  estimate from all eight Rif<sup>R</sup> isolates and the wild-type control is  $6.6 \times 10^{-8}$  per generation. This value is close to the reciprocal of the effective population size for this experiment ( $1/N_{\text{eff}} = 8 \times 10^{-8} \text{ cell}^{-1}$ ) after correcting for the fluctuations in

reported here relative to the fitness of the reference strain to allow direct comparisons of fitness changes with  $d$  and  $\Sigma$ . We found no evidence of nontransitive fitness interactions that would invalidate this normalization procedure; that is, the fitness of each evolved Rif<sup>R</sup> clone relative to its ancestor agreed with that predicted from its fitness measured relative to the reference strain and the fitness of its ancestor relative to the reference strain. (b) The effective rate at which the first successful beneficial mutations occur in each test strain ( $\mu$ ) plotted against that strain's initial fitness defect ( $d$ ). Values of  $\mu$  were inferred from the marker divergence analysis (see fig. 1). Error bars are 95% confidence intervals. There is no significant trend in  $\mu$ . The dashed line represents the average value over all strains. (c) The average fitness increase ( $\Sigma$ ) for each test strain over the full 640-generation evolution experiment plotted against that strain's initial fitness defect ( $d$ ). Single competitions of each clone isolated at 640 generations, one from each independently evolved replicate ( $n = 12$ ), against their oppositely marked  $rpoB$  mutant ancestors were used to measure  $\Sigma$ . These values were normalized relative to the reference strain's fitness to allow direct comparisons of fitness changes with  $d$  and  $s$ . Error bars are 95% confidence limits. A separate set of competitions between evolved clones and the Rif<sup>R</sup> reference strain that was used to establish transitivity gave essentially identical results. Trend lines are as described in (a). In all cases, this overall improvement in fitness is at least as great as the inferred single-step value, and most test strains appear to have reached at least the fitness level of the reference strain by this time.



**Fig. 3.** Fitness measurements of individual clones isolated near the end of the first selective sweep agree with estimates from the marker divergence analysis. We isolated six clones from each of three representative Rif<sup>R</sup>-8 populations at generation 128 and measured their fitness. The marker trajectories to this point are depicted on the left, and each bar on the right represents a clone isolated from the indicated Rif<sup>R</sup>-8 population. Gray and white bars denote the fitness of Ara<sup>-</sup> and Ara<sup>+</sup> isolates, respectively, measured relative to the reciprocally marked Rif<sup>R</sup>-8 ancestor strain in competition assays. Error bars are 95% confidence intervals estimated from replicate competitions ( $n = 6$ ). Based on the marker divergence trajectory analysis, the effective selective advantage for the first single-step beneficial mutations to fix in the Rif<sup>R</sup>-8 background is estimated as 0.30. For replicate populations #1 and #3, where one color dominates, the average fitnesses of the six evolved clones are 1.28 and 1.34, respectively. These values are thus in close agreement with the marker divergence estimate, and 10 of the 12 Rif<sup>R</sup>-8 populations show marker dynamics similar to these examples (see fig. 1). In contrast, population #2 is one of two Rif<sup>R</sup>-8 populations where the marker ratio trajectory diverged much more slowly. This difference appears to reflect, in part, clonal interference from beneficial mutations in the opposite marker background because an Ara<sup>-</sup> (red) clone isolated from this population has a fitness of  $1.12 \pm 0.05$  (95% confidence interval) relative to the ancestor. This population also appears to have not discovered mutations as beneficial as those found in most other populations, given that the average relative fitness of all Ara<sup>+</sup> (white) clones is only 1.19 at generation 128. Overall, these measurements of evolved isolates are consistent with the conclusions of the marker divergence analysis regarding the average sizes of the first beneficial mutations to sweep in the Rif-8 background.

population size during each daily cycle of dilution and re-growth (Patwa and Wahl 2008). This invariance in the estimated value of  $\mu$  is expected if the first beneficial mutations to fix occurred during the initial outgrowth phase from the small number of cells used to found each replicate population and if the distribution of beneficial mutations is smooth, such that sampling more mutants always results in finding one with a higher fitness, at least up to the population size of the experiment.

We found no evidence of reversion mutations or loss of rifampicin resistance during the evolution experiment. In

many cases, mutation to the ancestral *rpoB* sequence would have had a greater selective advantage than the first beneficial mutations that were successful, but reversion mutations were evidently too rare to be observed under our conditions (Levin et al. 2000). Indeed, the average point mutation rate per base pair in the ancestral strain (Sniegowski et al. 1997; Barrick et al. 2009) is more than two orders of magnitude lower than the average  $\mu$  for beneficial mutations in these experiments, suggesting that any given beneficial mutation that we observed in one of our experiments was drawn from a pool of hundreds of potential beneficial mutations of roughly similar advantage.

We also examined the total fitness increase of each test strain over all 640 generations of the evolution experiment by isolating a single evolved genotype with the predominant marker state from each final population and performing head-to-head competitions between these clones and their Rif<sup>R</sup> ancestors (fig. 2c). It is common for multiple beneficial mutations to arise in competing lineages before any one mutation completes a selective sweep in large asexual populations (Fogle et al. 2008; Barrick and Lenski 2009), and reversals in marker trajectories sometimes occur when more fit multiple mutants in one marker state surpass the single mutants in the other marker state that were responsible for the initial divergence (fig. 1). Genotypes that dominate each population by 640 generations have typically accumulated multiple beneficial mutations, and success at this point is dictated by integrating evolutionary dynamics over paths in the fitness landscape. The overall trend in this sustained multistep evolvability is very similar to that for single-step evolvability. Once again, there is a significantly positive slope to the regression ( $P = 0.00008$ ), and a strain's initial fitness defect is highly correlated with its evolvability ( $r^2 = 0.86$ ). The slope is still less than one ( $P = 0.001$ ), indicating that the difference in fitness between evolved Rif<sup>R</sup> and evolved reference strains persists, although it has decreased to 43%, on average, of the difference between their ancestors.

It is not surprising, perhaps, that lower fitness *E. coli* strains with *rpoB* mutations are more evolvable than their progenitors. Classical descriptions of peak climbing, including Fisher's geometric model, formalize this intuitive expectation (Orr 1998), and many studies of microbes have found rapid compensatory adaptation after deleterious mutations (Lenski 1988; Burch and Chao 1999; Levin et al. 2000; Moore et al. 2000; Reynolds 2000). Our experiments are unique, however, in how they systematically test a range of fitness defects and quantify evolvability on multiple timescales. In particular, we show that marker divergence experiments usefully summarize complex mutational neighborhoods in terms of effective parameters that reflect the abilities of different strains to adapt in a single step, given a specific environment and population structure.

The precision with which the fitness cost of each *rpoB* mutation predicts its effect on evolvability is striking. To the extent that the fitness landscapes on which biological organisms evolve have general properties and are not dominated by idiosyncratic interactions between mutations,



the simple relationship found here may thus serve as a useful null model for future experiments. Nonetheless, many questions remain. Do other deleterious mutations cause similar proportional increases in evolutionary potential? Do mutations in certain genes make microorganisms substantially more or less evolvable than expected from their fitness effects under this null model? Do beneficial mutations that are fixed during long-term adaptation to an environment invariably reduce evolvability, or are more evolvable genetic architectures favored even on microevolutionary timescales?

## Supplementary Material

Supplementary figures S1 and S2 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

## Acknowledgments

We thank N. Hajela, M. Adawe, and A. Melfi for laboratory assistance; N. Shores for detailed feedback concerning the analysis of marker divergence trajectories; M. Desai, J. Meyer, and R. Woods for helpful discussions; two anonymous reviewers for useful suggestions; and the High Performance Computing Center at Michigan State University. This research was supported by the National Institutes of Health (GM-087550 to J.E.B.); the National Science Foundation (DEB-0515729 to J.E.B.); the Michigan State University Frank Peabody Microbiology Student Research Fund and College of Natural Science Larry D. Fowler Endowment (to M.R.K.); and the Defense Advanced Research Projects Agency (HR0011-05-1-0057 and HR0011-09-1-0055 to R.E.L.).

## References

- Ahrens JH, Dieter U. 1982. Computer generation of Poisson deviates from modified normal distributions. *ACM Trans Math Software*. 8:163–179.
- Barrick JE, Lenski RE. 2009. Genome-wide mutational diversity in an evolving population of *Escherichia coli*. *Cold Spring Harbor Symp Quant Biol*. Advance Access published September 23, 2009, doi:10.1101/sqb.2009.74.018.
- Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243–1247.
- Bebbington AC. 1975. A simple method of drawing a sample without replacement. *Appl Statist*. 24:136.
- Blount ZD, Borland CZ, Lenski RE. 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci U S A*. 105:7899–7906.
- Burch CL, Chao L. 1999. Evolution by small steps and rugged landscapes in the RNA virus  $\phi$ 6. *Genetics* 151:921–927.
- Burch CL, Chao L. 2000. Evolvability of an RNA virus is determined by its mutational neighbourhood. *Nature* 406:625–628.
- Cao Y, Gillespie DT, Petzold LR. 2006. Efficient step size selection for the tau-leaping simulation method. *J Chem Phys*. 124:044109.
- Chao L, Cox EC. 1983. Competition between high and low mutating strains of *Escherichia coli*. *Evolution* 37:125–134.
- Colegrave N, Collins S. 2008. Experimental evolution: experimental evolution and evolvability. *Heredity* 100:464–470.
- de Visser JAGM, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE. 1999. Diminishing returns from mutation supply rate in asexual populations. *Science* 283:404–406.
- Draghi J, Wagner GP. 2008. Evolution of evolvability in a developmental model. *Evolution* 62:301–315.
- Earl DJ, Deem MW. 2004. Evolvability is a selectable trait. *Proc Natl Acad Sci U S A*. 101:11531–11536.
- Fasano G, Franceschini A. 1987. A multidimensional version of the Kolmogorov-Smirnov test. *Mon Not R Astron Soc*. 225:155–170.
- Fogle CA, Nagle JL, Desai MM. 2008. Clonal interference, multiple mutations and adaptation in large asexual populations. *Genetics* 180:2163–2173.
- Gerrish PJ, Lenski RE. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica*. 102/103:127–144.
- Hegreness M, Shores N, Hartl D, Kishony R. 2006. An equivalence principle for the incorporation of favorable mutations in asexual populations. *Science* 311:1615–1617.
- Herring CD, Glasner JD, Blattner FR. 2003. Gene replacement without selection: regulated suppression of amber mutations in *Escherichia coli*. *Gene* 311:153–163.
- Jeong H, Barbe V, Lee CH, et al. (11 co-authors). 2009. Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *J Mol Biol*. 394:644–652.
- Jin DJ, Zhou YN. 1996. Mutational analysis of structure-function relationship of RNA polymerase in *Escherichia coli*. *Methods Enzymol*. 273:300–319.
- Jolicoeur P. 1968. Interval estimation of the slope of the major axis of a bivariate normal distribution in the case of a small sample. *Biometrics* 24:679–682.
- Kassen R, Bataillon T. 2006. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nat Genet*. 38:484–488.
- Kirschner M, Gerhart J. 1998. Evolvability. *Proc Natl Acad Sci U S A*. 95:8420–8427.
- Lenski RE. 1988. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. II. Compensation for maladaptive pleiotropic effects associated with resistance to virus T4. *Evolution* 42:433–440.
- Lenski RE, Rose MR, Simpson SC, Tadler SC. 1991. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am Nat*. 138:1315–1341.
- Leroi AM. 2000. The scale independence of evolution. *Evol Dev*. 2:67–77.
- Levin BR, Perrot V, Walker N. 2000. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics* 154:985–997.
- MacLean RC, Buckling A. 2009. The distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa*. *PLoS Genet*. 5:e1000406.
- Matsumoto M, Nishimura T. 1998. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans Model Comput Sim*. 8:3–30.
- McBride RC, Ogbunugafor CB, Turner PE. 2008. Robustness promotes evolvability of thermotolerance in an RNA virus. *BMC Evol Biol*. 8:231.
- Moore FB-G, Rozen DE, Lenski RE. 2000. Pervasive compensatory adaptation in *Escherichia coli*. *Proc R Soc Lond B*. 267:515–522.
- Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* 52:935–949.
- Paquin CE, Adams J. 1983. Relative fitness can decrease in evolving asexual populations of *S. cerevisiae*. *Nature* 306:368–370.
- Patwa Z, Wahl LM. 2008. The fixation probability of beneficial mutations. *J R Soc Interface*. 5:1279–1289.
- Perfeito L, Fernandes L, Mota C, Gordo I. 2007. Adaptive mutations in bacteria: high rate and small effects. *Science* 317:813–815.

- Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 2007. Numerical recipes: the art of scientific computing. Cambridge: Cambridge University Press.
- R Development Core Team. 2009. R: a language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing. [cited 2009 July 1]. Available from: <http://www.R-project.org>.
- Renshaw E. 1993. Modelling biological populations in space and time. Cambridge: Cambridge University Press.
- Reynolds MG. 2000. Compensatory evolution in rifampin-resistant *Escherichia coli*. *Genetics* 156:1471–1481.
- Rozen DE, de Visser JA, Gerrish PJ. 2002. Fitness effects of fixed beneficial mutations in microbial populations. *Curr Biol.* 12:1040–1045.
- Sniegowski PD, Gerrish PJ, Lenski RE. 1997. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* 387:703–705.
- Sniegowski PD, Murphy HA. 2006. Evolvability. *Curr Biol.* 16:R831–R834.
- Wagner GP, Altenberg L. 1996. Perspective: complex adaptations and the evolution of evolvability. *Evolution* 50:967–976.
- Wahl LM, Gerrish PJ. 2001. The probability that beneficial mutations are lost in populations with periodic bottlenecks. *Evolution* 55:2606–2610.