

The Influence of Hitchhiking and Deleterious Mutation Upon Asexual Mutation Rates

Michael E. Palmer¹ and Marc Lipsitch

Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115

Manuscript received August 10, 2005

Accepted for publication February 15, 2006

ABSTRACT

The question of how natural selection affects asexual mutation rates has been considered since the 1930s, yet our understanding continues to deepen. The distribution of mutation rates observed in natural bacteria remains unexplained. It is well known that environmental constancy can favor minimal mutation rates. In contrast, environmental fluctuation (*e.g.*, at period T) can create indirect selective pressure for stronger mutators: genes modifying mutation rate may “hitchhike” to greater frequency along with environmentally favored mutations they produce. This article extends a well-known model of Leigh to consider fitness genes with multiple mutable sites (call the number of such sites α). The phenotypic effect of such a gene is enabled if all sites are in a certain state and disabled otherwise. The effects of multiple deleterious loci are also included (call the number of such loci γ). The analysis calculates the indirect selective effects experienced by a gene inducing various mutation rates for given values of α , γ , and T . Finite-population simulations validate these results and let us examine the interaction of drift with hitchhiking selection. We close by commenting on the importance of other factors, such as spatiotemporal variation, and on the origin of variation in mutation rates.

THE question of how natural selection affects asexual mutation rates has been considered since as early as 1930 (FISHER 1930), yet understanding is still not complete.

Most nonneutral mutations are unconditionally deleterious. In asexuals, this tends to penalize alleles that cause elevated genomic mutation rates, because they are tightly linked to the deleterious mutations they cause. Theory tells us that a population that is not adapting will favor mutation rates approaching zero (LIBERMAN and FELDMAN 1986).

In contrast, in an adapting population, nonzero mutation rates can be favored (LEIGH 1970, 1973; GILLESPIE 1981; ISHII *et al.* 1989). If the environment and genotype are aligned such that genotypes with higher fitness are genetically “nearby” (*e.g.*, a single mutation away), then alleles that increase mutation rate will tend to produce more of these beneficial mutations and “hitchhike” along with them to greater multiplicity. Laboratory studies have indeed shown that high mutation rates can be thus selected in an adapting population (COX and GIBSON 1974; CHAO and COX 1983; MAO *et al.* 1997; SNEGOWSKI *et al.* 1997). Studies of natural isolates have found that “mutator” strains, whose mutation rates are 10–1000 times the mutation rates of most members of their species, are found at measurable frequencies (*e.g.*, 1–14%) in many populations (JYSSUM 1960; GROSS and

SIEGEL 1981; LECLERC *et al.* 1996; MATIC *et al.* 1997). However, we lack a complete understanding of the dynamics that produce the observed frequency distributions of various mutator strengths.

Analysis by LEIGH (1970) and ISHII *et al.* (1989) describes how hitchhiking in an alternating environment dictates the evolutionarily stable strategy (ESS) mutation rate in a simple two-locus model. One dimorphic locus experiences direct selection conditioned on the environment, which alternates between two states. A second locus affects the mutation rate of the first locus, but itself experiences no direct selection. Maynard Smith (MAYNARD SMITH and PRICE 1973) added to the model of Leigh an ensemble of potential unconditionally deleterious mutations (call the number of such potential mutations γ).

In this article, we extend these analyses to consider an environmental locus comprising many mutable sites (call the number of sites α), to more realistically model a gene that is disabled in many configurations, but operational in only a single configuration. We combine the analytical result regarding a multisite environmental locus with the result of Maynard Smith, obtaining a single expression for the effects of a multisite environmental locus, together with an ensemble of potentially deleterious loci. This combination is significantly more realistic than Leigh’s original model if the biological system of interest is, for example, a bacterium that must gain or lose functionality of a certain gene to adapt to one, or the other, of a pair of alternating environments,

¹Corresponding author: 3841 18th St., San Francisco, CA 94114.
E-mail: genetics@mepalmer.net

while simultaneously avoiding deleterious mutations. Such an adaptive challenge might be faced, for example, by human pathogens, as they hop from host to host.

In addition to the above analytical results, which apply to infinite populations, this article also describes computer simulations of finite populations. These finite-population simulations include both a multisite environmental locus and multiple potentially deleterious loci, allowing us to observe the interaction of drift with the indirect selective effects described above. Under certain conditions that violate the assumptions of our analytical model, drift overwhelms the indirect selective effects upon the mutator gene deriving from hitchhiking with the environmental locus. The prevalent mutation rates in such scenarios tend to be considerably lower than the analytically derived ESS mutation rates, because the effects of hitchhiking become vanishingly weak relative to mutational load and drift.

Much recent theoretical and simulation work has considered observed bacterial mutation rates in light of three forces: (1) hitchhiking with beneficial mutations (selecting for higher mutation rate); (2) mutational load (selecting for lower mutation rate); and (3) the cost of fidelity (selecting for higher mutation rate). The importance of these three forces is actively debated. The cost of fidelity is especially poorly understood—in particular, how exactly it might increase as the mutation rate decreases, in the context of extant natural mechanisms of DNA replication.

In this article, we demonstrate that, for biologically reasonable values of α and γ , mutation rates in the observed range for bacteria may be attained without the need to invoke a “cost of fidelity” (*i.e.*, any direct selection on the mutator alleles themselves). This supports the possibility that selection due to the cost of fidelity is negligibly weak in the natural context, compared to the other two forces. Some recent articles have agreed (JOHNSON 1999) and some have disagreed (SNIEGOWSKI *et al.* 2000) with this suggestion.

LEIGH'S (1970) two-locus model essentially considers a balance of two indirect selective effects: (1) higher mutators are favored immediately after an environmental alternation, by indirect selection for higher mutation rates due to increased association with the newly favored allele; however, (2) lower mutators are favored during the wait to the next alternation, by indirect selection for lower mutation rates due to mutational load. The winning ESS mutation rate is a result of the competition between these two factors over the long term. (For an example of such a competition, please see supplemental plots SUPP1 and SUPP2 at <http://www.genetics.org/supplemental/>.)

Leigh indicated that, for his analysis to hold, T must be long enough to allow a complete substitution of the environmental allele at each environmental alternation. ISHII *et al.* (1989) show, more precisely, that complete allele substitution occurs for large values of the product

sT , rather than for large T alone. Ishii further shows that $u_{\text{ESS}}T$ is in fact not constant but depends on sT . For high sT , Ishii agrees with Leigh that $u_{\text{ESS}}T = 1$; however, as sT decreases below a certain threshold value, $u_{\text{ESS}}T$ gradually increases to a plateau of 1.6, in the case of periodic environmental alternation. For the case of random environmental alternation, Ishii finds that $u_{\text{ESS}}T = 1$ for high sT ; but that $u_{\text{ESS}}T$ decreases to zero for low sT (see Figure 2 of ISHII *et al.* 1989). In other words, the system obeys the simple rule $u_{\text{ESS}}T = 1$ when sT is high; otherwise the behavior is more complex because complete substitution of the environmental allele is not guaranteed at each environmental alternation.

The impact of drift: Both Leigh and Ishii considered infinite populations. Neither author remarked on the vanishing strength of indirect selection upon the mutator alleles as sT approaches zero and the completeness of allele substitution decreases. In an infinite population, and given an infinite number of repeated environmental alternations, the absolute strength of the indirect selection does not matter (only the relative strengths among the various mutator strains matter) in determining the ESS winning mutation rate. In finite populations, however, drift is present; therefore, as the absolute strength of the indirect selection decreases at vanishing sT , it is eventually overwhelmed by drift. In this article, simulations are used to determine the minimum value of sT required to produce indirect selection on the mutator alleles due to hitchhiking, in the presence of drift.

RESULTS

Analytical results: Modeling multiple mutable sites at the environmental locus: Leigh's simple environmental locus mutates from one allele to the other and back again at symmetric rates. In contrast, natural genes typically contain multiple mutable sites at which a mutation will cause the gene to be disabled. To restore functionality, the particular site that suffered an adverse mutation must experience a reverse mutation. (We neglect the possibility of compensatory mutation.) Therefore, the rates at which the gene is disabled, and reenabled, are not symmetric. Furthermore, once the gene is disabled by a single adverse mutation, it may continue to suffer further adverse mutations; the chance of an exact reversal of multiple adverse mutations then becomes slim. We wish to incorporate these effects into a model of hitchhiking. The derivation below borrows the general reasoning of LEIGH (1970).

Consider an environmental gene represented by multiple sites that have the potential to experience adverse mutation; one of more such mutations will disable the gene and its resulting fitness effect. Let α be the number of such sites. Each site holds one of two states, labeled by the digits 1 (one) or 0 (zero). If all sites are in state 1, then the entire gene is in state A ; however,

if one or more of the sites are in state 0, then the entire gene is in state a . All sites independently mutate between states 1 and 0 at the symmetric rate u . As in Leigh's model, the environment alternates periodically between favoring state A and state a , at period T ; fitnesses are $1 + s$ in the favored environment and 1 otherwise.

Consider the moment just before an alternation from environment A to a . The population is nearly all-1 (all sites in state 1, meaning that the environmental gene is overall in state A). Due to immediately recent mutation, there is a fraction αu that are single-0 (a single site, of α total sites, is in state 0; the gene overall is in state a). Immediately after the alternation, all the all-1 individuals are quickly eliminated and the single-0 mutants survive. Over the T generations to the next alternation, these single-0 individuals lose a fraction αu per generation: u per generation are lost back to all-1 and quickly decrease, and $(\alpha - 1)u$ are transformed to double-0 (the double-0 mutants do not immediately die, but are "doomed" to die at the next alternation), for a total of αu lost per generation. At the moment of the next environmental alternation back to favoring A , there will be only a fraction u of the single-0 mutants that immediately switch to all-1 and survive. The double-0 mutants (and the smaller number of triple-0 mutants, etc.) cannot make the double (triple, etc.) switch to all-1 and quickly decrease. Over the next T generations, the all-1 mutants lose αu per generation to the single-0 state, which is now disfavored. Overall, following Leigh's method, this indicates that we must maximize (over a complete environmental cycle of two epochs $A \rightarrow a \rightarrow A$):

$$\alpha u e^{-\alpha u T} u e^{-\alpha u T} = \alpha u^2 e^{-2\alpha u T}. \quad (1)$$

Maximizing over u produces the following result for an environmental gene with multiple mutable sites capable of disabling the gene:

$$u_{\text{ess}} T = 1/\alpha. \quad (2)$$

The foregoing argument assumes that the population at the time of each environmental alternation is nearly pure and that a single generation of mutants seeds the population that takes over after the alternation; in fact, mutation–selection balance may be approached within each environmental epoch. This adds a factor of s_c^{-2} to Equation 1 but does not change the optimal value of u .

The addition of multiple unconditionally deleterious sites: Maynard Smith (MAYNARD SMITH and PRICE 1973) made a similar extension to Leigh's model, incorporating multiple potential sites for unconditionally (not dependent on the environment) deleterious mutations. If one calls the number of such sites γ , Maynard Smith maximized

$$u e^{-\gamma u T} \quad (3)$$

to obtain

$$u_{\text{ess}} T = 1/\gamma. \quad (4)$$

Combined result: Maynard Smith's result can be combined with that of Equation 2, to yield the following u_{ess} mutation rate for a genome with both a multisite environmental locus (α -sites) and multiple potential unconditionally deleterious sites (γ -sites). Maximizing the following over u ,

$$\alpha u e^{-(\alpha+\gamma)uT} u e^{-(\alpha+\gamma)uT} = \alpha u^2 e^{-2(\alpha+\gamma)uT}, \quad (5)$$

produces the following combined result:

$$u_{\text{ess}} T = 1/(\alpha + \gamma). \quad (6)$$

This result holds both for infinite and (approximately) for finite populations, provided that sT is high enough to generate strong allele reversal at each environmental alternation. Below, we show by simulation of finite populations that, at low sT , drift overwhelms the indirect selective effects leading to this ESS mutation rate.

Simulation results: We have created a population genetics simulator for finite populations of organisms with genomes of several loci, such as the two-locus genome defined by Leigh. The simulator resembles several simulators used recently to study mutation rates (TADDEI *et al.* 1997; TENAILLON *et al.* 1999, 2000; TRAVIS and TRAVIS 2002, 2004). It employs discrete generations and additive fitness, and it cycles all individuals through the steps of (1) differential reproduction, (2) sampling down to the carrying capacity, and (3) mutation. The simulator explicitly models a population of individuals with genomes of multiple loci, each locus potentially containing multiple mutable sites.

Genome definition: Our simulated genomes contain two classes of loci: fitness loci, which directly affect fitness, and mutator loci, which affect global mutation rate.

Fitness loci directly affect fitness and may be of two types: *AND fitness* loci and *additive fitness* loci. Table 1 contains one fitness locus of each type. The fitness of each genotype is initially set to the base fitness value indicated for each environment (A or a). Both types of fitness loci then contribute additively to this fitness. An AND fitness locus can be used to model a single gene that is operational only in one configuration of all its sites (all sites in state 1) and disabled in any other configuration (one or more of its sites in state 0). In Table 1, if all of the α -sites of the AND fitness locus are in the 1 state, it contributes (once) an additive fitness of s_c (or $-s_c$ in environment a ; this is the source of the dependence of fitness on the environment); otherwise, if one or more sites are in the 0 state, the entire locus is ignored. In contrast, the additive fitness locus in Table 1 models an ensemble of γ potential unconditionally deleterious mutations. It additively contributes the value $-s_d$ to the fitness of the genotype for each of its γ -sites

TABLE 1
A three-locus model

Locus type	Site count	Strength	Wild-type forward rate	Wild-type reverse rate
Environment A: base fitness = 1.0, base mutation factor = m_b				
AND fitness	α	s_c	u	u
Additive fitness	γ	$-s_d$	u	u
Multiplicative mutator	m_k	m_r	0	0
Environment a: base fitness = $1.0 + s_c$ base mutation factor = m_b				
AND fitness	α	$-s_c$	u	u
Additive fitness	γ	$-s_d$	u	u
Multiplicative mutator	m_k	m_r	0	0

The genome contains three loci. The first locus produces environmentally sensitive selection and can be disabled by a mutation at any of α sites. The second locus allows for γ potential deleterious mutations, which have no environmental dependence. The third locus controls the mutation rate of the sites in the other two loci: it will range from um_b to $um_b m_r^{m_k}$.

that are in the 1 state for that genotype; sites in the 0 state are ignored.

Mutator loci affect mutation rates, but not fitness. The genome in Table 1 has one multiplicative mutator locus, which contains m_k dimorphic (states 1 or 0) sites. The state of the sites in this locus for a particular genotype determines the “mutation factor” of that genotype. The value of m_b , the mutation factor for genotype i , is computed as $m_i = m_b m_r^i$, where m_b is the base mutation factor, m_r is the multiplicative strength of each mutator site, and i indicates the number of mutator sites in the 1 state for that particular genotype (sites in the 0 state are ignored). The number i ranges from zero to m_k for the genome in Table 1. The mutation factor of a genotype multiplicatively increases the specified “wild-type” mutation rates for all sites in that genotype (the per site wild-type mutation rate of the fitness loci in Table 1 is u and is symmetric); hence, mutator strains experience multiplied mutation rates at all sites, compared to the wild-type strain. In Table 1, the sites in the mutator locus do not mutate: their mutation rates are zero.

Estimating u_{ess} at many points in parameter space: The simulator described above can perform a competition of multiple mutator strains, typically resulting in dominance by one strain of a particular mutation rate. To precisely estimate the optimal mutation factor, and hence the ESS per site mutation rate, for a particular genome and a particular set of parameters (*e.g.*, T , s_c , s_d , α , γ), we conduct multiple competitions of increasing resolution of the mutation factor (by decreasing the value m_r from Table 1 and appropriately varying m_b). This allows convergence to a precise estimate of the ESS mutation rate u_{ess} .

We created a total of 2400 (T , s_c , s_d , α , γ) five-dimensional tuples (or “parameter sets”), by crossing five values of T (64, 256, 1024, 4096, and 16,384) by four values of s_c (6.25×10^{-4} , 2.5×10^{-3} , 0.01, and 0.04) by four

values of s_d (1.5625×10^{-4} , 6.25×10^{-4} , 2.5×10^{-3} , and 0.01) by five values of α (1, 5, 10, 50, and 100) by six values of gamma (0, 1, 5, 10, 50, and 100).

For each of these 2400 tuples, we applied a recursive process (of up to three levels of recursion) to converge to an estimate u_{ess} at that point in the five-dimensional parameter space. (Please see supplemental plot SUPP3 at <http://www.genetics.org/supplemental/> for an illustration of this process.) At each level of the convergence process for a particular tuple, we caused 18 mutator strains to compete. The strains used at the first level are separated from each other by a large mutation factor (*i.e.*, m_r is initially large); we “zoom in” for increasing precision as follows. The competition at a particular level for a particular tuple is repeated several times in distinct runs of the simulator, each with a different random seed. We declare a particular strain to be the winner of a run if it has been the most populous strain for a number of generations equal to the maximum of 20,000 and $10T$. In addition, runs are terminated at 500,000 generations even if there is no winner. We require that at least four runs, within a maximum of 10 attempts, obtain a winning strain. If a particular level yields a tight enough distribution of winners (as specified below), then we zoom in for increasing precision: we start another competition of 18 new mutator strains by dividing the width of 2 steps from the previous level into 18 smaller steps, centering them around the average winning value from the previous level.

Periodic environmental alternation: We first present simulation results using periodic environmental alternation at constant period T .

Overview of tuples that converged exceptionally: The four values of s_c and the five values of T we used yielded eight distinct $s_c T$ values (0.04, 0.16, 0.64, 2.56, 10.24, 40.96, 163.84, and 655.35). Figure 1 plots, by $s_c T$, the disposition of each of the 2400 tuples after convergence to an estimate of u_{ess} was attempted. The total number of

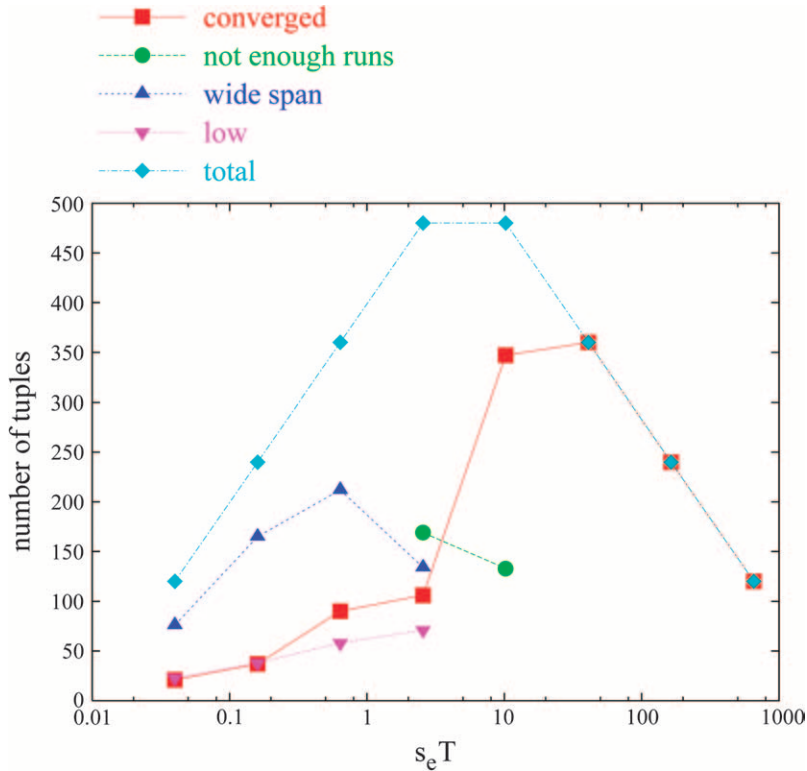


FIGURE 1.—Final disposition of all tuples under periodic alternation. The total number of tuples attempted at each $s_e T$ value is indicated by the line marked “total.” The count of well-converged tuples at each $s_e T$ value is indicated by the line marked “converged.” Tuples not converging but finishing with various exceptional conditions are also indicated.

tuples at each $s_e T$ value is shown by the line marked “total.” The line marked “converged” indicates the count of well-converged tuples at each $s_e T$ value. Well-converged tuples are those whose convergence process resulted in an estimate of u_{ess} (to some precision level), rather than terminating with an exceptional condition. Most tuples with $s_e T$ values of ≥ 10.24 were well converged, because the indirect selection resulting from hitchhiking is quite strong at high $s_e T$. However, many tuples of lower $s_e T$ terminated with an exceptional condition, yielding no estimate of u_{ess} for that tuple.

One such exceptional condition is indicated in Figure 1 by the line labeled “not enough runs.” At a particular level of the convergence process, if we do not obtain four runs (within a maximum of 10 attempts) with a declared winner, that level of convergence terminates with the exceptional condition of not enough runs. If this occurs at the first level of the convergence process for that tuple, the entire tuple receives the not enough runs exception; we can make no estimate of u_{ess} for that tuple. If it occurs at a subsequent level (of a maximum of three levels), the convergence process is terminated, and the tuple receives a valid $u_{\text{ess}} T$ value with appropriately larger error bars. Generally, this exception indicates that two or more strains were alternately dominating the population, often with one preferentially associated with environmental state A and another with state a . In Figure 1, this can be seen to sometimes occur for intermediate values of $s_e T$ (2.56 and 10.24).

Occurrence of another exceptional condition is plotted by the line labeled “wide span.” It can happen

that we do obtain a winner in at least four runs of a particular competition, but that those winners are spaced too widely apart (*i.e.*, the lowest and highest winning strains were separated by more than two steps), indicating that it would be fruitless to seek a more precise estimate of u_{ess} . If this occurs at the first level of the convergence process, then we do not return any u_{ess} value at all and terminate with the exceptional condition of wide span; however, if it happens at a subsequent level, we do return a u_{ess} value, including appropriately larger error bars. As shown in Figure 1, this generally occurs at low values of $s_e T$, indicating that drift is playing a large role in the determination of the winning u_{ess} at these values.

Another exceptional termination condition is labeled “low”; this indicates that, at the first level of the convergence process, the lowest of the strains was the winner for at least one run. In this case, it is possible that our preset range of mutation rates did not go low enough to contain the true winning u_{ess} value; therefore we cannot continue the convergence process normally. Again, if this occurs past the first level of convergence, we do return a u_{ess} value and give it appropriately larger error bars. However, if this happens at the first level, we terminate exceptionally with the low condition. At the first level of the convergence process for each of the 2400 tuples, m_b was 0.0019531, m_r was 4, and m_k was 17. Since a base mutation rate of $u = 7.45058 \times 10^{-9}$ was used, this created 18 strains with per site mutation rates of from 1.45519×10^{-11} to 0.25 (the latter being an absurdly high mutation rate for natural organisms),

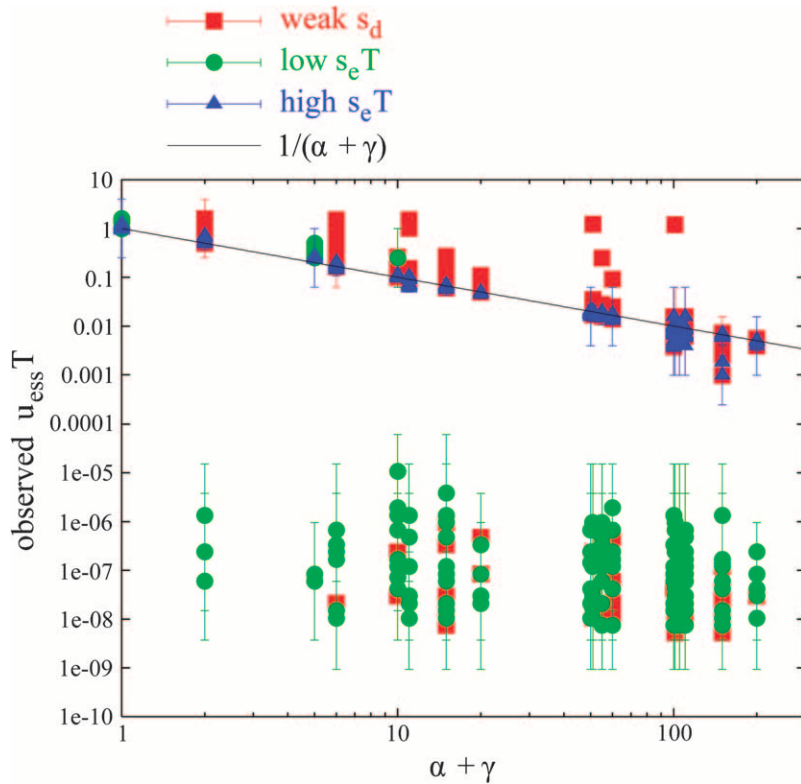


FIGURE 2.—Overview of all well-converged tuples for periodic environmental alternation. The tuples are divided into three classes: “high $s_e T$ ” points, which generally obeyed Equation 6; “weak s_d ” points, for which potential deleterious mutations are present but weak; and “low $s_e T$ ” points, for which hitchhiking due to selection becomes vanishingly weak. The three classes of points are further analyzed in Figures 3–6. The black line indicates the prediction of Equation 6.

separated by factors of 4. The low condition therefore indicates that a strain with per site mutation rate of 1.45519×10^{-11} won at least once, at the first level of convergence, for a given tuple.

Other exceptional termination conditions included “high” and “both high and low,” which are self-explanatory; however, these did not occur for any of the 2400 tuples.

Overview of well-converged tuples: We now consider the 1321 of 2400 tuples that converged normally, yielding an estimate of u_{ess} to some level of precision. Figure 2 plots these 1321 points. For each tuple, the empirical estimate of u_{ess} multiplied by T is shown on the y -axis; the x -axis indicates the value of $\alpha + \gamma$ for the respective tuples. A black line plots the value of $u_{\text{ess}} T = 1/(\alpha + \gamma)$ as predicted by Equation 6. Error bars about each point indicate how closely the convergence process was able to estimate $u_{\text{ess}} T$ for that tuple. (More precisely, the error bars indicate the highest and lowest $u_{\text{ess}} T$ values of the four winning strains of the four trials at the lowest level of convergence attained by that tuple. Note that the presence of a tuple in Figure 2 implies that it was well converged.) Three separate classes of points are plotted:

1. “Weak s_d ” points are those with $s_d/s_e < 0.1$, as well as $\gamma > 0$ (*i.e.*, the genome did contain one or more potentially deleterious sites). The remaining tuples (those with stronger s_d or no deleterious sites at all, *i.e.*, $\gamma = 0$) are split into two further groups.
2. Points with $s_e T < 10.24$ are classified as “low $s_e T$.” This was subjectively determined to be the threshold

below which most tuples did not adhere to the prediction of Equation 6.

3. Finally, the remaining points with $s_e T \geq 10.24$ are labeled “high $s_e T$.” The high $s_e T$ points generally match the prediction of Equation 6. The next several paragraphs discuss these three classes of points in detail; they illuminate some details of how a finite population diverges from the prediction of Equation 6 at the borderline between sufficiently and insufficiently complete substitution of the environmental allele at each environmental alternation. (*i.e.*, in the range of $s_e T = 10.24$).

High $s_e T$ tuples: Of the 1321 well-converged tuples, 611 fell into the category “high $s_e T$ ” (*i.e.*, $s_e T \geq 10.24$, and not falling into the “weak s_d ” category, below). These are the points that are expected to match the prediction of Equation 6. Figure 3 plots the high $s_e T$ points, separated by $s_e T$ value. Of these 611 points, 496 (81%) matched the prediction of Equation 6 (*i.e.*, their error bars span the black line). It can be seen from Figure 3 that the points that strayed furthest from the prediction were generally those with the relatively low $s_e T = 10.24$. Points with high $(\alpha + \gamma)$ are particularly prone to diverge when $s_e T$ is low.

Weak s_d tuples: Of the 1321 well-converged tuples, 512 fell into the category of weak s_d (*i.e.*, $s_d/s_e < 0.1$ and $\gamma > 0$). These are points that have at least one ($\gamma > 0$) potentially deleterious site, but a low s_d , relative to s_e . Figure 4 plots these points, separated by $s_e T$ value. Some

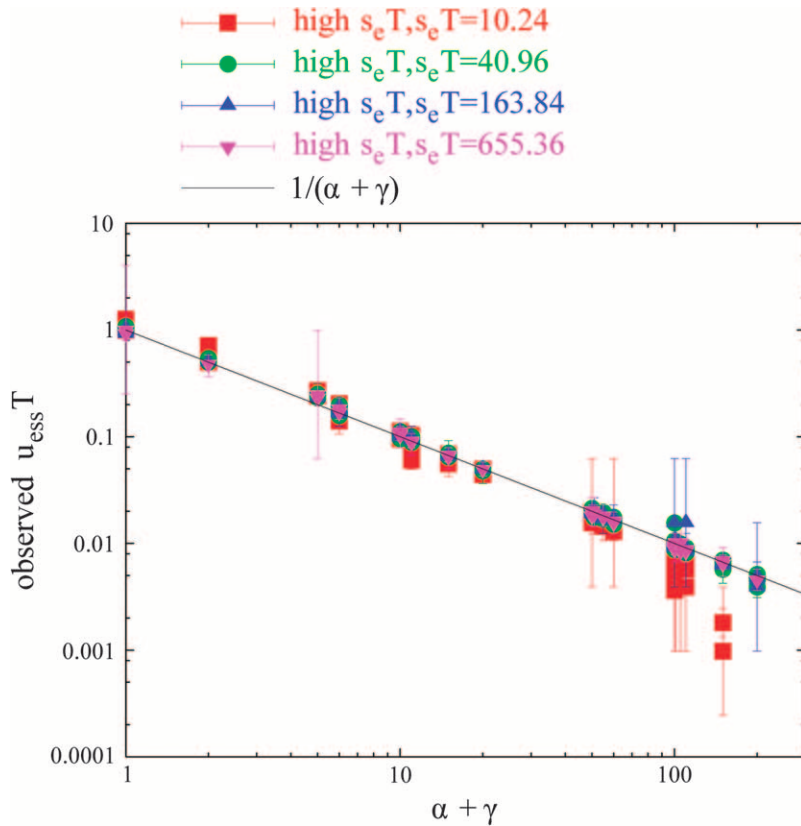


FIGURE 3.—High $s_e T$ tuples for periodic alternation. The “high $s_e T$ ” tuples from Figure 2 are plotted, separated by $s_e T$ value. Those of lower $s_e T$ value ($s_e T = 10.24$) tend to be the ones that stray from the prediction of Equation 6.

of these weak s_d points, particularly those with moderate $s_e T$ ($s_e T = 10.24$ or 40.96) or low $s_e T$, behave as if there are no potentially deleterious mutations ($\gamma = 0$) and nearly obey the relationship $u_{\text{ess}} T = 1/\alpha$. For example, there are a number of points in a horizontal line around $u_{\text{ess}} T = 1.2$; for these points, $\alpha = 1$ (α -value is not indicated on the plot). (We believe that the equality to 1.2 instead of 1.0 is congruent with the observation of ISHII *et al.* (1989) that $u_{\text{ess}} T$ should gradually increase from 1.0 to 1.6 as $s_e T$ decreased through intermediate values such as 10.24.) Similarly, there are other points in a horizontal line at $u_{\text{ess}} T \approx 0.2 = \frac{1}{5}$, with α -values of 5 (α -value is not indicated on the plot), and points in a horizontal line at $u_{\text{ess}} T \approx 0.1 = \frac{1}{10}$, with α -values of 10. There are also some tuples with both weak s_d and low $s_e T$ (Figure 4, red squares), which attain very low $u_{\text{ess}} T$ values—see the next paragraph.

Low $s_e T$ tuples: Of the 1321 well-converged tuples, 198 fell into the category “low $s_e T$ ” ($s_e T < 10.24$, and not falling into the weak s_d category, above). Figure 5 plots these points separately by $s_e T$ value. In general, the pattern is evident that adherence to Equation 6 drops off as $s_e T$ decreases, particularly for higher values of $(\alpha + \gamma)$. Points of lower $s_e T$ and/or higher $(\alpha + \gamma)$ all obtain low $u_{\text{ess}} T$ values. Recall that Figure 5 plots only points that converged well; referring back to Figure 1, we note that, for tuples of low $s_e T$, many more points terminated with the exceptional conditions of low or wide span than converged well. This yields the overall

picture that indirect selection due to hitchhiking with the environmental allele for tuples of low $s_e T$ becomes negligibly weak: their low value of $u_{\text{ess}} T$ is due to selection against mutational load and/or drift.

Random (geometrically distributed) environmental alternation: We also performed the convergence process for the same 2400 tuples under geometrically distributed random environmental alternation at average period T . The behavior of the 2400 tuples, as compared to the periodic case, differed in two important ways: first, we would place the subjective threshold between low and high $s_e T$ at 40.96, as compared to 10.24 for periodic alternation. Second, for points of lower $s_e T$, there is not an abrupt transition from strong hitchhiking to no hitchhiking: a moderately strong hitchhiking effect can exist, producing moderate values of $u_{\text{ess}} T$. Figure 6 shows the low $s_e T$ points in the random alternation case. (For space reasons, we do not show all the various plots for the random case. However, see supplemental plots SUPP4–SUPP8 at <http://www.genetics.org/supplemental/>, which plot the tuples in the random case in the same manner that Figures 1–5 plot the tuples in the periodic case.) If we compare Figure 6 to Figure 5 (the corresponding plot of low $s_e T$ points for the periodic case), we see these novel intermediate values of $u_{\text{ess}} T$. Since the population is no longer subjected to a single period of environmental alternation, but rather to a geometric distribution of periods, we may expect strong hitchhiking during some epochs

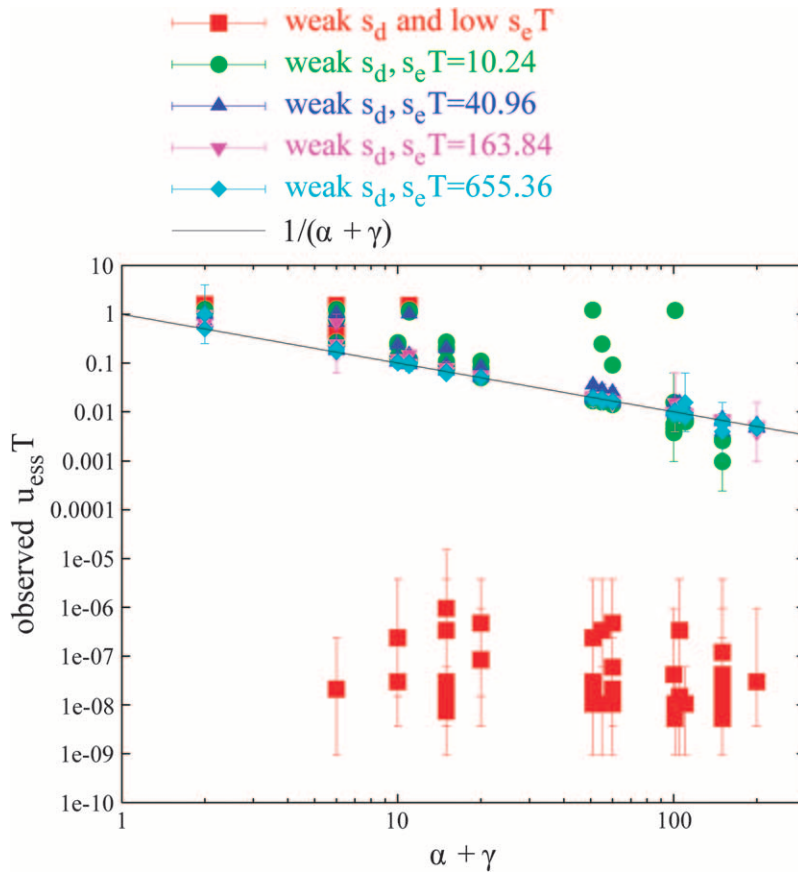


FIGURE 4.—Weak s_d tuples for periodic alternation. The “weak s_d ” tuples (those for which deleterious mutations are present, but weak) from Figure 2 are plotted, separated by $s_e T$ value. Those for which s_d is weak and $s_e T$ is intermediate ($s_e T = 10.24$) or low stray most frequently from the prediction of Equation 6. Weak s_d tuples for which $s_e T$ is intermediate often behave as if the deleterious locus is being ignored ($\gamma = 0$), appearing to approach the rule $u_{\text{ess}} T = 1/\alpha$; see text.

and no hitchhiking in others. The ratio of long-enough to short-enough epochs should gradually decrease with decreasing $s_e T$. In Figure 6, it is apparent that the higher $s_e T$ tuples more closely approach adherence to Equation 6 than the lower ones. Moreover, although we do not plot them here, the high $s_e T$ points ($s_e T \geq 40.96$) in the stochastic case largely (302 out of 370 tuples, or 82%) did adhere to Equation 6.

In summary: If $s_e T$ is high, producing strong hitchhiking, then $u_{\text{ess}} T = 1/(\alpha + \gamma)$ (Equation 6) is a good predictor of $u_{\text{ess}} T$. As s_d becomes weaker, deleterious mutations begin to be ignored in the hitchhiking dynamic. As $s_e T$ becomes weaker, the hitchhiking effect itself becomes so weak as to be overwhelmed by drift (with or without potentially deleterious mutations), and $u_{\text{ess}} T$ is instead determined by selection against mutational load (if potentially deleterious mutations are present) and drift.

DISCUSSION

Relation to previous work: The findings described here are consistent with the findings of previous theoretical studies and provide a context for generalizing some of them. For example, TADDEI *et al.* (1997) and TANAKA *et al.* (2003) found that mutators could be selected for during periods of environmental change

(*i.e.*, at times when beneficial mutations were available) but might then disappear after all such mutations had been found. Our model, building on the work of LEIGH (1970), ISHII *et al.* (1989), and MAYNARD SMITH and PRICE (1973), places the findings of Taddei and Tanaka in the context of a longer time span including many episodes of environmental change and stasis. Moreover, by abstracting from the distributions of deleterious and lethal mutations of Taddei *et al.*, we find by simulation that the rate of mutations at least approximately one-tenth as deleterious as the beneficial environmental mutation can be incorporated into the ESS rate by using the formula of Maynard Smith. We also clarify the relationship of drift to hitchhiking selection, which was not considered by Leigh, Ishii, and Maynard Smith.

Our findings also provide a context for the findings of TRAVIS and TRAVIS (2002), who found that mutator strains are most favored in regimes of intermediate environmental alternation frequency. In our model, this finding naturally arises because excessively rapid alternation (low T) for a given s_e produces “low $s_e T$,” making hitchhiking ineffective and leading to selection for lower mutation rates. Alternately, very slow alternation (high T) also leads to selection for reduced mutation rates due to increased back and deleterious mutations. This framework allows extrapolation of Travis and Travis’ simulation findings; for example, we predict that

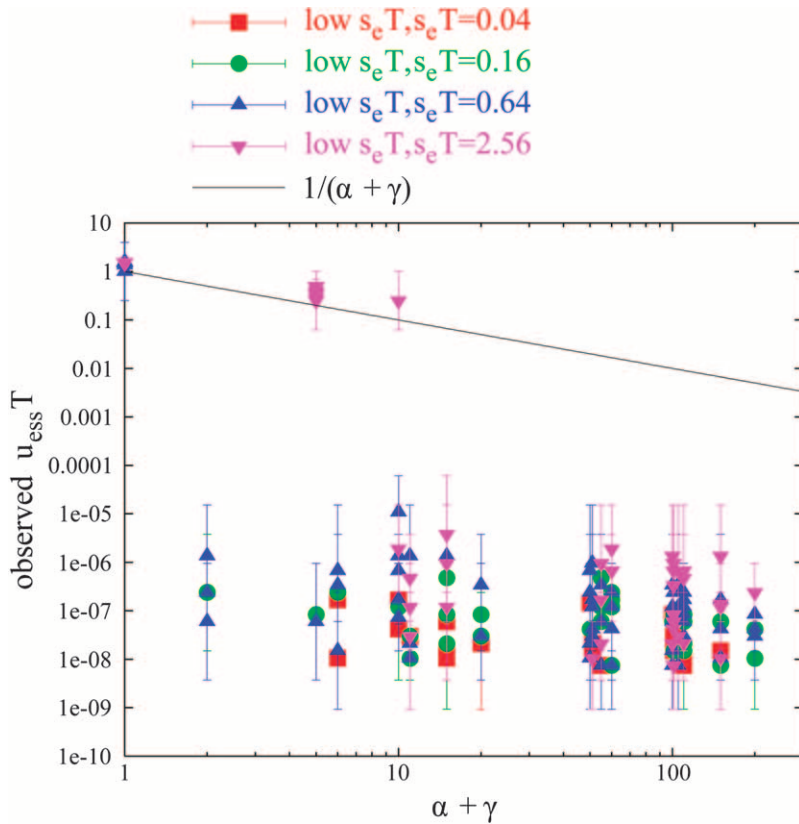


FIGURE 5.—Low $s_e T$ tuples under periodic alternation. The “low $s_e T$ ” tuples from Figure 2 are plotted, separated by $s_e T$ value. Generally, adherence to Equation 6 decreases as $s_e T$ decreases and/or $(\alpha + \gamma)$ increases.

as the selective value of the environmental locus (s_e) declines, environmental alternation periods (T) that were previously long enough to support a given mutator will become too short.

Comparison to observed natural mutation rates: Observed baseline mutation rates in bacteria are low: DRAKE *et al.* (1998) estimate the per base pair, per generation, mutation rate to be $u_{bp} = 5.4 \times 10^{-10}$ in *Escherichia coli* “well adapted to laboratory conditions.” A long-term laboratory study by Lenski found a slightly lower estimate of $u_{bp} = 1.44 \times 10^{-10}$ for nonmutator populations (LENSKI *et al.* 2003; LENSKI 2004).

The simple model of LEIGH (1970) would require a $T > 10^9$ to produce such rates; it clearly does not apply—1 billion generations is an unrealistically long wait between selective sweeps of beneficial mutations. Does Equation 6 attain more reasonable rates, given natural values of α , γ , and T ?

Rough estimates of the values of γ and α in *E. coli* are as follows. KIBOTA and LYNCH (1996) also consider *E. coli* in an artificially stable laboratory environment. They estimate the genomic rate of deleterious mutation to be $u_g = 2 \times 10^{-4}$, as a lower bound. (The average deleterious fitness effect of the mutations they measured was 0.012 per mutation.) Since $\gamma \times u_{bp} = u_g$ (where u_{bp} is the mutation rate per base pair of 5.4×10^{-10}), we arrive at an estimate of $\gamma \sim 370,000$ (as a lower bound). As for α , we assume that the disabling or the reenabling of a single protein is what is required to

adapt to two alternating environments. If the length of the *E. coli* genome is 4.6×10^6 bp, and *E. coli* produces ~ 4290 proteins, then an upper bound on the number of sites that could disable the average protein is ~ 1000 . Therefore we estimate α at 1000—a small fraction of γ .

Using these estimates of α and γ , a T of ~ 5000 generations would be required to attain the low baseline mutation rates (5.4×10^{-10}) found by DRAKE *et al.* (1998). It has been suggested that the turnover of *E. coli* populations in a human host occurs on a timescale of weeks to a month (CAUGANT *et al.* 1981). Similar carriage durations are reported for such bacteria as *Streptococcus pneumoniae* (EKDAHL *et al.* 1997) and *Haemophilus influenzae* (FADEN *et al.* 1995), while *Staphylococcus aureus* (SCANVIC *et al.* 2001) and *Neisseria meningitidis* (DE WALS and BOUCKAERT 1985) have durations of the order of 5–10 months. While 1 month may represent ~ 1500 bacterial generations in laboratory medium, one plausible estimate of the net doubling time for *E. coli* in a human host is ~ 40 hr (SAVAGEAU 1989), so that 1 month may represent as few as 20 generations. Thus, it seems safe to say that many bacteria change human hosts with a timescale of 10^1 – 10^4 generations, only the very upper range of which approaches the order of 5000 generations predicted by Equation 6. Therefore, it appears that Equation 6 does not easily explain the low rates found by Lenski and Drake, at least for the potential source of repeated environmental variation represented by host-to-host transfer (however, see

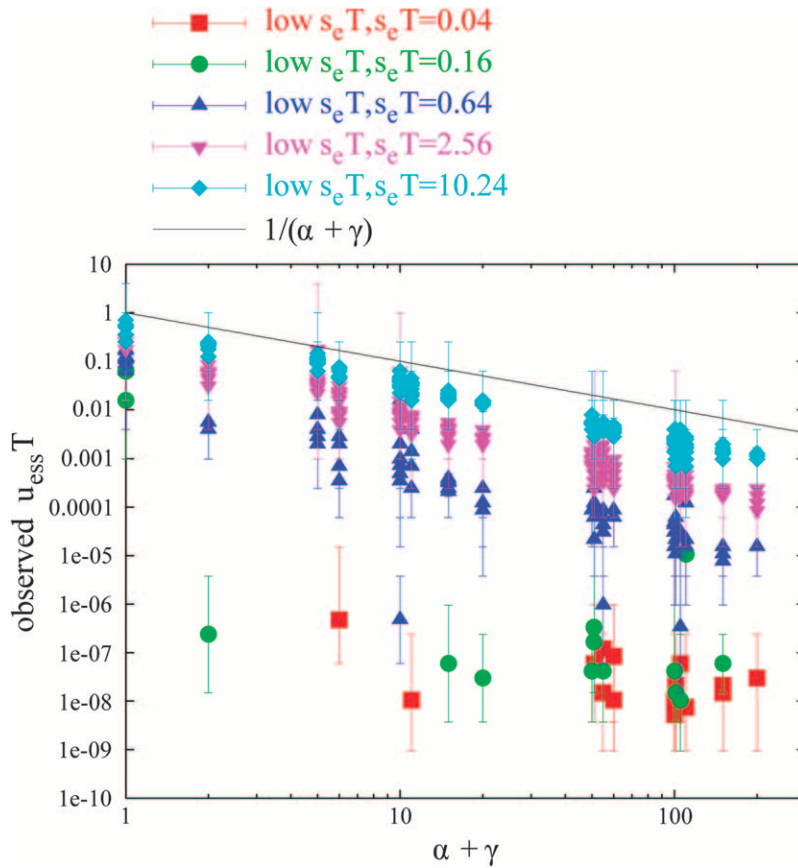


FIGURE 6.—Low $s_e T$ tuples under random environmental alternation. In the random alternation case, intermediate values of $u_{\text{ess}} T$ appear for intermediate values of $s_e T$, indicating a gradual transition from the presence to the absence of hitchhiking. Compare low $s_e T$ tuples under periodic alternation in Figure 5.

the next section, *What combination of forces produces the low baseline mutation rate?*

What about bacteria “in the wild,” *e.g.*, living within and between human hosts? These strains might be expected to experience significantly more fluctuating selection than strains maintained in the laboratory. MATIC *et al.* (1997) collected 504 isolates of commensal and pathogenic *E. coli* from human hosts. They screened the strains for forward mutagenesis in the *lacI* gene, finding that nearly 90% of these strains were nonpapillating (generally indicating lower mutation rate—*i.e.*, these are the nonmutators), having an average mutation rate to rifampicin resistance of 1×10^{-8} /individual/generation, with low variance. Previously, JIN *et al.* (1988) counted all point mutations producing resistance to rifampicin, finding a total of 17. If we divide the rate of mutation to rifampicin resistance according to Matic by this number, we obtain a per base pair mutation rate of 5.9×10^{-10} , approximately equal to that found by Drake. Therefore, although they were collected “from the wild,” 90% of all strains tested by Matic appear to maintain low rates in the range of those of Drake and Lenski. Matic found that the other 10% of the strains did produce papillation (tending to indicate higher mutation rate). These strains produced an average mutation rate to rifampicin resistance of 2.6×10^{-7} , corresponding to a rate of 1.5×10^{-8} /bp (compa-

table to the 100-fold mutators found by Lenski). What do we make of Matic’s observations? Apparently, 90% of strains in the wild demonstrate the same mutation rate as strains that were intentionally maintained in a stable environment. There appear to be common features that pertain to both “stable” and “unstable” environments: the same baseline rate and typical rates for common mutators (*i.e.*, the 100 \times mutator strains).

Two questions stand out.

What combination of forces produces the low baseline mutation rate? KIMURA (1967) introduced a hypothesized cost of fidelity to explain why bacterial mutation rates do not decrease to zero. The cost of fidelity would exert an upward force on mutation rates. Kimura and others, *e.g.*, SNIĘGOWSKI *et al.* (2000), have assumed that this force was in balance with the opposing downward force of mutational load. Instead, our model suggests that a balance between hitchhiking selection and mutational load could be responsible for setting the optimum; in other words, the invocation of a cost of fidelity may not be required. However as we note above, something seems to be depressing observed mutation rates lower than our model predicts (given our estimates of α , γ , and T and assuming strong hitchhiking under high $s_e T$). Therefore, in addition to the problem of explaining why mutation rates are nonzero, it is also difficult to explain why they are as low as they are.

Remaining within the confines of our model for the moment, for Equation 6 to produce lower rates, either γ or T would have to increase (since α cannot become much larger relative to γ). Taking a $T \sim 500$ generations, in the middle range of plausible estimates given above, and α and γ as above, Equation 6 yields u_{ess} of 5.4×10^{-9} — ~ 10 times higher than Drake's rate. Since the number of base pairs in *E. coli* is only $\sim 4.6 \times 10^6$, it would be unlikely that γ could increase 10 times from our previous estimate of 370,000.

As for T , we do note that TRAVIS and TRAVIS (2004), investigating spatial variation using a two-patch model in which the environmental alternation of the patches was at the same frequency but out of phase, found that this provided a sort of "refuge" for higher mutators, due to some individuals migrating between the patches and effectively experiencing a higher T . Whether such effects could produce the missing 10-fold difference is beyond our scope here. Nonetheless, the complexity observed in TRAVIS and TRAVIS' (2004) study of selection on mutation rate with only two populations suggests that a variety of interesting phenomena may appear in such models. Bacteria in human hosts may be highly compartmentalized. Much work remains to be done to understand how the present model can be generalized to understand the evolution of mutation rates in such metapopulations.

But what if the model's assumption of high $s_e T$ were relaxed? First of all, let us assume that the random alternation case, rather than the periodic case, better resembles the distribution of T between selective sweeps of adaptive mutations experienced by bacteria. Recall from Figure 6 that certain tuples in the random alternation case do strongly converge to $u_{\text{ess}} T$ values that are 10- and 100-fold less than that predicted by Equation 6. These are the higher of the values in the low $s_e T$ category; therefore we call them "intermediate $s_e T$." For these cases, selection on the environmental allele is not quite strong enough to produce a complete substitution of alleles within the duration of many epochs. If the intermediate $s_e T$ case applies, then the low rates observed by Drake could be produced within the confines of our current model (with the high $s_e T$ assumption relaxed).

Why does environment seem not to affect the low baseline mutation rate? The second question that arises from Matic's observed distribution of mutation rates is: Why do bacteria in the wild (presumably a strongly fluctuating environment) maintain the same low baseline rate found in the laboratory? The simplest explanation is that Drake's and Lenski's strains were not in the laboratory long enough for a new ESS mutation rate to dominate; they simply reflect the wild ESS rate. For example, Lenski's populations were well documented for 20,000 generations at the time of publication in 2003 (LENSKI *et al.* 2003), but what happened to the inoculating strain prior to this is less clear. It had been in

the laboratory since at least 1966, but, for example, might have been frozen for much of that time. In our simulations, it was easy to construct cases in which an ESS mutation rate would not fix for tens or hundreds of thousands of generations—for example, when T is very long, as would be expected in a stable environment.

Genetic accessibility of mutation rates: Another important influence upon the distributions of mutation rates found in nature could be the genetic accessibility of various mutator strengths. Obviously, the theoretical ESS mutation rate cannot come to dominate a population if no strain with that rate arises in the first place. Consider that only a subset of the genes in the *E. coli* genome have an influence on mutation rate. All possible point mutations to these genes produce a finite set of distinct mutator strains (and most such mutations simply produce unviable strains). One might imagine a graph in which the nodes represent the common mutator strains of *E. coli* genetically near to a wild type, and the edge weights represent the transition (*e.g.*, mutation) probabilities among the strains. This graph may constitute a hidden genetic context for the evolution of mutators. When a beneficial mutation is available (*i.e.*, within a single point mutation or so, due to the confluence of environmental and genetic factors), the population may temporarily crowd into higher mutation rate strains via hitchhiking. However, when beneficial mutations are not available, the population will tend to retreat into nearby lower mutation rate strains. (One article making a simulation study of this pursuit of mutation–selection balance was TADDEI *et al.* 1997, which allowed forward and reverse mutation between a wild type and a single mutator type.) The indirect selection felt by the various mutator strains may well be described by Equations 5 and 6, but any individual does not have complete flexibility as to what mutation rates its progeny may adopt next; these are dictated by the graph, which evolves only over the very long term. Several mutational "hops" back and forth through relatively favored and disfavored mutation rates may be required to attain the optimum rate. Such hopping, perhaps in the face of drift, may take time—and moreover any "optimum" rate may be ephemeral. Therefore, the distribution of mutator strengths observed in natural bacteria probably cannot be fully explained by considering only environmental influence, without any genetic influence. Much work remains to be done in this area.

Interplay between mutation rate and population size: Finally, we mention a subtlety of the model that depends on effective population size. In our derivation of Equations 5 and 6, we assumed that each mutator strain had an ample number of both the favored and disfavored alleles, meaning that exponential growth could begin immediately after the environmental alternation to a new favored allele. Therefore, we neglected any effect due to a delay in "discovery" of the favored allele by each

strain or delay due to its struggle against loss by drift. (Recall that Leigh's original model considered infinite populations, so this was not a concern.) Such delays would tend to decrease the effectiveness of hitchhiking at low mutation rates relative to the reciprocal of the effective population size. In addition, they would also decrease the penalty due to deleterious mutations. Further discussion is beyond our scope, but this effect is discussed in TANAKA *et al.* (2003).

In summary, we believe that we have furthered the understanding of the indirect selection experienced by mutators due to hitchhiking and deleterious mutation. A full and detailed understanding of the distribution of mutation rates observed in bacteria, however, may require the consideration of genetic constraints as well, and the nature of these constraints is not currently well understood. In addition, we believe that further complication will be uncovered by models of temporally and spatially complex environments.

We thank Richard Moxon and Christopher Bayliss, of Oxford, for helpful discussions throughout the course of this work. We thank several anonymous reviewers of the manuscript in progress for important comments and suggestions. M.P. thanks his wife for her support of this work. This work was supported by National Institutes of Health grant 1R21AI055825, to principal investigator Marc Lipsitch.

LITERATURE CITED

- CAUGANT, D. A., B. R. LEVIN and R. K. SELANDER, 1981 Genetic diversity and temporal variation in the *E. coli* population of a human host. *Genetics* **98**: 467–490.
- CHAO, L., and E. C. COX, 1983 Competition between high and low mutating strains of *Escherichia coli*. *Evolution* **37**: 125–134.
- COX, E. C., and T. C. GIBSON, 1974 Selection for high mutation rates in chemostats. *Genetics* **77**: 169–184.
- DE WALS, P., and A. BOUCKAERT, 1985 Methods for estimating the duration of bacterial carriage. *Int. J. Epidemiol.* **14**: 628–634.
- DRAKE, J. W., B. CHARLESWORTH, D. CHARLESWORTH and J. F. CROW, 1998 Rates of spontaneous mutation. *Genetics* **148**: 1667–1686.
- EKDAHL, K., I. AHLINDER, H. B. HANSSON, E. MELANDER, S. MOLSTAD *et al.*, 1997 Duration of nasopharyngeal carriage of penicillin-resistant *Streptococcus pneumoniae*: experiences from the South Swedish Pneumococcal Intervention Project. *Clin. Infect. Dis.* **25**: 1113–1117.
- FADEN, H., L. DUFFY, A. WILLIAMS, D. A. KRYSIOFIK and J. WOLF, 1995 Epidemiology of nasopharyngeal colonization with non-typeable *Haemophilus influenzae* in the first 2 years of life. *J. Infect. Dis.* **172**: 132–135.
- FISHER, R. A., 1930 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- GILLESPIE, J. H., 1981 Mutation modification in a random environment. *Evolution* **35**: 468–476.
- GROSS, M. D., and E. C. SEGEL, 1981 Incidence of mutator strains in *Escherichia coli* and coliforms in nature. *Mutat. Res.* **91**: 107–110.
- ISHII, K., H. MATSUDA, Y. IWASA and A. SASAKI, 1989 Evolutionarily stable mutation rate in a periodically changing environment. *Genetics* **121**: 163–174.
- JIN, D. J., W. A. WALTER and C. A. GROSS, 1988 Characterization of the termination phenotypes of rifampicin-resistant mutants. *J. Mol. Biol.* **202**: 245–253.
- JOHNSON, T., 1999 Beneficial mutations, hitchhiking and the evolution of mutation rates in sexual populations. *Genetics* **151**: 1621–1631.
- JYSSUM, K., 1960 Observations on two types of genetic instability in *Escherichia coli*. *Acta Pathol. Microbiol. Scand.* **48**: 113–120.
- KIBOTA, T. T., and M. LYNCH, 1996 Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. *Nature* **381**: 694–696.
- KIMURA, M., 1967 On the evolutionary adjustment of spontaneous mutation rates. *Genet. Res.* **9**: 23–34.
- LECLERC, J. E., B. LI, W. L. PAYNE and T. A. CEBULA, 1996 High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. *Science* **274**: 1208–1211.
- LEIGH, JR., E. G., 1970 Natural selection and mutability. *Am. Nat.* **104**: 301–305.
- LEIGH, JR., E. G., 1973 The evolution of mutation rates. *Genetics* **73**: 1–18.
- LENSKI, R. E., 2004 Phenotypic and genomic evolution during a 20,000-generation experiment with the bacterium *Escherichia coli*. *Plant Breed. Rev.* **24**: 225–265.
- LENSKI, R. E., C. L. WINKWORTH and M. A. RILEY, 2003 Rates of DNA sequence evolution in experimental populations of *Escherichia coli* during 20,000 generations. *J. Mol. Evol.* **56**: 498–508.
- LIBERMAN, U., and M. W. FELDMAN, 1986 Modifiers of mutation rate: a general reduction principle. *Theor. Popul. Biol.* **30**: 125–142.
- MAO, E. F., L. LANE, J. LEE and J. H. MILLER, 1997 Proliferation of mutators in a cell population. *J. Bacteriol.* **179**: 417–422.
- MATIC, I., M. RADMAN, F. TADDEI, B. PICARD, C. DOIT *et al.*, 1997 Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science* **277**: 1833–1834.
- MAYNARD SMITH, J., and G. R. PRICE, 1973 The logic of animal conflict. *Nature* **246**: 15–18.
- SAVAGEAU, M., 1989 *Escherichia coli* habitats, cell types, and molecular mechanisms of gene control. *Am. Nat.* **122**: 732–744.
- SCANVIC, A., L. DENIC, S. GAILLON, P. GIRY, A. ANDREMONTE *et al.*, 2001 Duration of colonization by methicillin-resistant *Staphylococcus aureus* after hospital discharge and risk factors for prolonged carriage. *Clin. Infect. Dis.* **32**: 1393–1398.
- SNIEGOWSKI, P. D., P. J. GERRISH and R. E. LENSKI, 1997 Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387**: 703–705.
- SNIEGOWSKI, P. D., P. J. GERRISH, T. JOHNSON and A. SHAVER, 2000 The evolution of mutation rates: separating causes from consequences. *BioEssays* **22**: 1057–1066.
- TADDEI, F., M. RADMAN, J. MAYNARD-SMITH, B. TOUPANCE, P. H. GOUYON *et al.*, 1997 Role of mutator alleles in adaptive evolution. *Nature* **387**: 700–702.
- TANAKA, M. M., C. T. BERGSTROM and B. R. LEVIN, 2003 The evolution of mutator genes in bacterial populations: the roles of environmental change and timing. *Genetics* **164**: 843–854.
- TENAILLON, O., B. TOUPANCE, H. LE NAGARD, F. TADDEI and B. GODELLE, 1999 Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. *Genetics* **152**: 485–493.
- TENAILLON, O., H. LE NAGARD, B. GODELLE and F. TADDEI, 2000 Mutators and sex in bacteria: conflict between adaptive strategies. *Proc. Natl. Acad. Sci. USA* **97**: 10465–10470.
- TRAVIS, E. R., and J. M. TRAVIS, 2004 Mutators in space: the dynamics of high-mutability clones in a two-patch model. *Genetics* **167**: 513–522.
- TRAVIS, J. M., and E. R. TRAVIS, 2002 Mutator dynamics in fluctuating environments. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **269**: 591–597.

Communicating editor: M. K. UYENOYAMA