Horizontal Acquisition of Divergent Chromosomal DNA in Bacteria: Effects of Mutator Phenotypes

Jeffrey P. Townsend,*,1 Kaare M. Nielsen,*,2 Daniel S. Fisher† and Daniel L. Hartl*

*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138 and
†Department of Physics, Harvard University, Cambridge, Massachusetts 02138

Manuscript received October 10, 2002
Accepted for publication January 25, 2003

ABSTRACT

We examine the potential beneficial effects of the expanded access to environmental DNA offered by mutators on the adaptive potential of bacterial populations. Using parameters from published studies of recombination in E. coli, we find that the presence of mutators has the potential to greatly enhance bacterial population adaptation when compared to populations without mutators. In one specific example, for which three specific amino acid substitutions are required for adaptation to occur in a 300-amino-acid protein, we found a 3500-fold increase in the rate of adaptation. The probability of a beneficial acquisition decreased if more amino acid changes, or integration of longer DNA fragments, were required for adaptation. The model also predicts that mutators are more likely than nonmutator phenotypes to acquire genetic variability from a more diverged set of donor bacteria. Bacterial populations harboring mutators in a sequence heterogeneous environment are predicted to acquire most of their DNA conferring adaptation in the range of 13–30% divergence, whereas nonmutator phenotypes become adapted after recombining with more homogeneous sequences of 7–21% divergence. We conclude that mutators can accelerate bacterial adaptation when desired genetic variability is present within DNA fragments of up to ~30% divergence.

SEQUENCING of bacterial genes (Guttman and Dykuizen 1994; Feng et al. 1997; Feil et al. 1999, 2000) and genomes (Rivera et al. 1998; Jain et al. 1999; Ochman et al. 2000) and the frequent observations of mosaic patterns in bacterial genes, including some encoding antibiotic resistance (Spratt et al. 1989; Downson et al. 1990; Spratt 1994; Claverys et al. 2000) and pathogenicity traits (Bessen and Hollingshead 1994; Kroll et al. 1998), strongly suggest that recombination with divergent DNA is of importance for bacterial evolution. Moreover, the number of observed nucleotide changes due to recombination in natural populations of bacteria has been found to range from approximately equal to the number of changes due to mutation in Bacillus spp. (Roberts and Cohan 1995) to up to 100-fold higher in Escherichia coli (Guttman and Dykuizen 1994), Neisseria spp. (Feil et al. 1999), and Streptococcus spp. (Feil et al. 2000). Thus, whereas the exact mechanism enabling the horizontal gene transfer often remains unidentified, it is clear that bacteria are exposed to foreign DNA molecules at appreciable frequencies under natural conditions. Furthermore, laboratory studies of the uptake of exogenous chromosomal DNA in bacteria (Bowler et al. 1994; Denamur et al. 2000) have also demonstrated that recombination can mediate the process of adaptive evolution.

A major barrier to bacterial acquisition of exogenous DNA is encoded by the methyl-directed DNA mismatch repair genes (e.g., mutS, mutL, and mutH in E. coli; Rayssiguet al. 1989) that efficiently reduce the rate of incorporation of divergent DNA into the bacterial genome (Matic 1995; Matic et al. 1996). Interestingly, bacteria with defects in their mut genes have been found in environmental, commensal, and pathogenic isolates at frequencies ranging from 1% (Leclerc et al. 1996; Matic et al. 1997) to 100% in multiple antibiotic-resistant pathogens (Oliver et al. 2000); these frequencies are well above that expected from the neutral mutation rate. These hypermutable phenotypes, called mutators, have been suggested to enhance adaptation of bacterial populations by increasing the supply rate of beneficial mutations during periods of strong selection (Taddei et al. 1997b; Tenaillon et al. 1999). An increase in frequency of mutators has also been observed in laboratory experiments where strong selection has been applied (Mao et al. 1997; Notley-McRobb et al. 2002).

Bacterial adaptation could be enhanced by horizontal genetic transfer when selection requires multiple mutations for adaptation to occur. Mutational processes may be insufficient for adaptation in such instances because the time required for a set of mutations to arise sequentially via neutral or deleterious intermediate stages is too long. Interestingly, many bacteria with the mutator phenotype also show an increased capability to recom
A divergent DNA that is inaccessible to the prevailing bacterial populations is unclear. Recombination with divergent DNA, for example, has been described for both wild-type and mutator strains of *E. coli* (Vulic et al. 1997), *Bacillus subtilis* (Zawadski et al. 1995; Majewski and Cohan 1998), and *Streptococcus pneumoniae* (Majewski et al. 2000) as

\[
\rho(v) = R_0 e^{-\omega v},
\]

where \( R_0 \) is the rate of recombination per generation at zero divergence, and \( \omega \) is the exponential rate of decrease of recombination with increasing sequence divergence.

The rate of recombination in bacterial populations, \( r \), as a function of sequence divergence, \( v \), has been described for both wild-type and mutator strains of *E. coli* (Vulic et al. 1997), *Bacillus subtilis* (Zawadski et al. 1995; Majewski and Cohan 1998), and *Streptococcus pneumoniae* (Majewski et al. 2000) as

\[
r(v) = R_0 e^{-\omega v},
\]

where \( R_0 \) is the rate of recombination per generation at zero divergence, and \( \omega \) is the exponential rate of decrease of recombination with increasing sequence divergence.

The link between recombinogenic mutators with small values of \( \omega \) (and thus increased access to interspecies genetic diversity) and the adaptive potential of bacterial populations is unclear. Recombination with divergent DNA is expected to cause deleterious effects, for instance, by introducing nucleotide changes resulting in amino acid substitutions that impair protein function and/or alter gene expression and regulation. Yet, in cases where several nucleotide changes are required for substantial adaptation to occur, mutators can potentially benefit the evolving population through acquiring divergent DNA that is inaccessible to the prevailing bacterial phenotypes.

To clarify to what extent mutators provide bacterial populations with beneficial DNA without simultaneously introducing deleterious effects, we have modeled the outcome of recombination events in *E. coli* populations of wild-type (nonmutators) and mutator phenotypes. We hypothesized that rare recombination events facilitated by recombinogenic mutator bacteria provide a source of variation inaccessible by point mutations and thus may play a key role in evolutionary innovation.

**MODEL**

We consider a single-gene model of homologous recombination to examine the limits to and the consequences of the acquisition of divergent but homologous chromosomal DNA by different bacterial phenotypes. Parameters and functions introduced by the model are defined in Tables 1 and 2. Realistic parameter values for Equation 1 were abstracted from empirical studies of recombination with divergent DNA in experimental *E. coli* populations (Vulic et al. 1997). They are also representative for gene transfer frequencies observed in many other bacterial species (see reviews by Davison 1999; Dro¨ge et al. 1999). Note, however, that there is considerable uncertainty about and probably considerable variation in the recombination parameter \( R_0 \), as this reflects not only integration into chromosomal DNA, but also the availability of exogenous DNA and its rate of uptake into the cell, both of which depend strongly on the environmental conditions (Nielsen et al. 1997, 2000).

The probability of a successful recombination event depends on the distribution of adaptive sequences in the bacterial environment. We therefore considered two representative but contrasting ecological habitats, where the relative concentration of DNA of increasing sequence divergence was given by a unimodal \( \beta \)-function

\[
e(\nu) = C_\beta \nu^\beta (1 - \nu)^\kappa,
\]

where \( C_\beta \) is a normalization constant, and \( \eta \) and \( \kappa \) are chosen to reflect a homogeneous \((\eta = 0, \kappa = 9)\) or heterogeneous environment \((\eta = 3, \kappa = 1)\) of the recipient bacterium (Figure 1). In a homogeneous environment, organisms with DNA sequences close to those of the bacterial recipient predominate. The elevated proportion of DNA with little sequence divergence allows a high rate of recombination, but also limits the

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range of study</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v )</td>
<td>0–1</td>
<td>Nucleotide divergence</td>
</tr>
<tr>
<td>( \omega )</td>
<td>16–203</td>
<td>Exponential rate of decrease of recombination with sequence divergence</td>
</tr>
<tr>
<td>( l )</td>
<td>0–1000</td>
<td>Length of coding DNA fragment, in codons</td>
</tr>
<tr>
<td>( n )</td>
<td>1–10</td>
<td>Number of amino acid changes required for a beneficial effect</td>
</tr>
<tr>
<td>( q )</td>
<td>0.05–0.067</td>
<td>Probability that a differing amino acid is the desired residue</td>
</tr>
<tr>
<td>( \eta )</td>
<td>0–3</td>
<td>First parameter for the beta distribution ( \eta(\nu) )</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>1–9</td>
<td>Second parameter for the beta distribution ( \kappa(\nu) )</td>
</tr>
<tr>
<td>( R_0 )</td>
<td>( 1.6 \times 10^{-7} - 5 \times 10^{-6} )</td>
<td>Rate of recombination at zero divergence</td>
</tr>
</tbody>
</table>
| \( C_\beta \) | 10–20        | Normalization constant for \( e(\nu) \), equal to \( \Gamma(\eta + \kappa + 2)/\Gamma(\eta + 1) \times \Gamma(\kappa + 1) \)
| \( A_\eta \) | 0–4            | Parameter for \( a(\nu) \) describing probability that diverged amino acids are deleterious |
| \( X_\kappa \) | 0–4           | Parameter for \( a(\nu) \) describing amino acid divergence |
probability of acquisition of more divergent DNA. In contrast, in the heterogeneous environment modeled, the recipient species contribute an insignificant fraction of the DNA available. Thus, recombination events will be rare, but those that do occur may result in the acquisition of substantially divergent DNA. Such rare events may provide a source of variation inaccessible to bacteria through mutational processes alone and may, thus, play a key role in bacterial evolution.

Most interspecies recombination events that occur in coding regions are expected to be harmful due to deleterious amino acid replacements. The probability of acquisition of a lethally deleterious gene fragment of amino acid length $l$ and divergence $v$ was modeled, on the basis of independence of changes of each codon, as

$$ f(v, l) = 1 - (1 - a(v)x(v))^l, \quad (3) $$

where $a(v)$ is the probability of amino acid replacement at a codon given nucleotide divergence $v$ and $x(v)$ is the probability of a lethally deleterious amino acid replacement acquired from an organism with DNA divergence $v$. The function $a(v)$ was obtained from distance matrices of naturally occurring bacteria by regression. The functional form used,

$$ a(v) = A_0v^2(1 - v) + \delta^2, \quad (4) $$

satisfies boundary conditions $a(v = 0) = 0$ and $a(v = 1) = 1$, so that no nucleotide divergence yields no amino acid divergence, and maximum nucleotide divergence yields maximum amino acid divergence. We used a value of $A_0 = 3$, obtained by least squares from 2.2-kb fragments of the bacterial $mutS$ gene from 27 isolates of the genus Acinetobacter (K. M. Nielsen, D. Young, N. O. Ornston and D. L. Hartl, unpublished data); preliminary analysis did not show elevated recombinational activity in the $mutS$ gene over other genes as had been reported for various E. coli isolates (Denamur et al. 2000; Brown et al. 2001). Moreover, other values of $A_0$ do not materially affect the results presented here.

The function $x(v)$, which parameterizes how often a novel amino acid substitution acquired through recombination with DNA of divergence $v$ will be lethally deleterious, was approximated at small amounts of divergence to be proportional to $v^2$. In the context of speciation, Orr (1995) has argued that a simple interpretation of the Dobzhansky-Muller model yields conflicts at a rate proportional to at least $a(v)^2$. This functional form reflects the probability that a single altered amino acid in one strain is incompatible with any of the altered amino acids in the other strain, a probability proportional to $a(v)$. Therefore we expect that $x(v) \approx a(v)$ for small $v$. Since we also expect $x(v = 0) = 0$ and $x(v = 1) = 1$, we chose a function of the same form as $a$:

$$ x(v) = X_0v^2(1 - v) + \delta^2. \quad (5) $$

Some experimental approaches to finding $x(v)$ have been developed (Coyne and Orr 1989), but we do not have data for E. coli from which it can be reliably extracted; for definiteness, $X_0 = 1$ is used in our examples. As we shall see, even in the heterogeneous model environment, acquisition of beneficial DNA is dominated by divergence, $v$, that is still relatively small. In the Appendix, we use small $v$ approximations to derive simple analytic approximations for the quantitative results from the models. Under the parameter ranges discussed, these approximations were within a factor of two or better. Thus, although $X_0$ is a relevant and undetermined parameter, other functional forms of

---

**TABLE 2**

<table>
<thead>
<tr>
<th>Function</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r(v)$</td>
<td>Rate of acquisition of DNA sequence by recombination</td>
</tr>
<tr>
<td>$c(v)dv$</td>
<td>Fraction of ambient DNA in interval $(v, v + dv)$</td>
</tr>
<tr>
<td>$f(v, l)$</td>
<td>Probability that an acquired fragment has a deleterious effect</td>
</tr>
<tr>
<td>$h(v, l, n)$</td>
<td>Probability that an acquired fragment has a beneficial effect</td>
</tr>
<tr>
<td>$p_d(v, l)dv$</td>
<td>Probability of a deleterious recombination event, within interval $(v, v + dv)$</td>
</tr>
<tr>
<td>$p_b(v, l)dv$</td>
<td>Probability of a beneficial recombination event, within interval $(v, v + dv)$</td>
</tr>
</tbody>
</table>

---

**Figure 1**—Probability density, $c(v)$, of accessible DNA as a function of sequence divergence in a model homogeneous or heterogeneous environment. The total concentrations of accessible DNA were equal in the two environments; the curves both integrate to unity.
The probability of a deleterious recombination event occurring in the DNA environment specified by \( c(v) \) is
\[
p_d(v, l) = f(v, l) r(v) c(v). \tag{6}
\]
Using this formulation, Figure 2A illustrates the probability that recombination events with DNA encoding a range of amino acid divergences, at three constant nucleotide divergence levels, are deleterious in a protein of average length \( l = 300 \) amino acids. For those same nucleotide divergence levels, how the probability that a recombination event is deleterious varies with increasing length, \( l \), of the recombining fragment is shown in Figure 2B. This reveals a strong relation between the length of DNA acquired and the likelihood of its deleteriousness \( \partial f(v, l) / \partial l \) is strongly positive], particularly at higher divergence levels. Zawadski and Cohan (1995) reported that the size of DNA segments integrated in \( B. subtilis \) transformation varied with the size of the initial donor molecule and decreased with...

\[ h(v, l, n) = q^r(a(v))^{s(1 - a(v))}^{r(1 - a(v)x(v))}^{c(n)}, \tag{7} \]
with \( q \) representing the probability that an amino acid change at one of the specified sites is of the required type. We used \( q = 1/20 \), which assumes equal usage of amino acids. Utilizing a matrix of amino acid abundance from \( E. coli \), one obtains \( q \approx 1/15 \) (Blattner et al. 1997). The fact that the most likely amino acid changes involve only one nucleotide change would also modify the appropriate \( q \). These effects are readily taken into account and do not significantly change the comparisons reported here. For simplicity and generality, we have assumed that the presence of one of the set of desired amino acids does not affect the probabilities of the presences of the other \( n - 1 \); because of their potential for having been of similar collective benefit to the donor organism, \( h(v, l, n) \) of the above form is likely an underestimate. The probability density for acquiring a specific beneficial set of amino acids from bacteria with divergence \( v \) in the given environment is hence
\[ p_b(v, l, n) = h(v, l, n) r(v) c(v). \tag{8} \]
The total probability of a beneficial recombination event in relation to the distribution of accessible sequences is calculated as \( \int f l p_b(v, l, n) dv \), and that of a deleterious recombination event as \( \int f l p_d(v, l, n) dv \). An analysis based on small \( v \) revealed the degree to which these results depend on the parameters \( \omega, l \), and \( n \) (see the Appendix); we also derived the most probable divergence from which beneficial acquisition may occur. For wild-type organisms, the most probable divergence is approximately
\[
ationalauthors
Recombination in Bacterial Mutator Populations

Figure 3.—Expected number of minimum lengths of sequence identity (MLSI), with 14-, 20-, or 25-bp segments of absolute sequence identity, in a 900-bp sequence. The dashed line indicates an expectation of one MLSI.

on the long-term evolution of the recombination rate. On the basis of the described parameters (Table 1), the probabilities of beneficial acquisition of divergent DNA from two contrasting environments were estimated for mutators and nonmutators of E. coli, assuming that three specific amino acid changes are required for adaptation to occur. This number is within the range of introduced amino acid replacements (n = 1–19) discovered in horizontally acquired gene fragments in Streptococcus (Dowson et al. 1989). It is a large enough number of substitutions to convey a selectable phenotype, yet small enough to demonstrate the importance of recombination even for the acquisition of only slightly changed DNA.

Employing parameters determined from recombination studies in E. coli, we compared two strains (Vulic et al. 1997): wild type (R0 = 0.05, ω = 64) and mutator mutSΔ (R0 = 0.05, ω = 16). Additionally, we considered a strain in which the wild-type mutS gene is abundantly expressed, resulting in yet lower recombination frequencies than those of the wild type (R0 = 0.0016, ω = 203). Several studies have indicated that the minimum length of sequence identity (MLSI) required for initiation of recombination in E. coli and B. subtilis is 20–25 bp (Watt et al. 1985; Shen and Huang 1986; Majewski and Cohan 1999). If two or more homologous regions are present, shorter stretches may suffice. Homologous recombination between DNA strands exceeding ~30–35% of randomly dispersed divergence will thus occur at negligible frequency due to the improbability of a MLSI to initiate the process. Figure 3 shows the expected number of MLSIs present in a 900-bp DNA fragment as a function of DNA divergence, estimated as (3l − m + 1) (1 − v)^e, where 3l is the length in nucleotides of the homologous sequence and m is the length of the MLSI. The probability that there is an MLSI is thus of order e−ωm, suggesting that the minimum value of ω is determined by the length, m, of a MLSI. The observed value of ω = 16 for a mutSΔ strain of E. coli may thus be dominated by the availability of MLSI. Substantially smaller values of ω are therefore unlikely to be found.

The calculated probability p0(v) of acquisition of n = 3 specific amino acids in a fragment of length l = 300 amino acids from a donor with divergence v (Zawadzki and Cohan 1995) is shown for each of the three strains in Figure 4. Both the heterogeneous and the homogeneous model environments are shown. The effects of mutators were most striking in the heterogeneous environment, where we found that the mutator was 3500-fold more efficient than nonmutators at acquiring beneficial sequences by recombination. A mutator was most likely (95% of the time) to acquire adaptive DNA sequences from other organisms in the range of 13–30% divergence (maximum at 21%), whereas nonmutator bacteria less efficiently accessed DNA and did so in the range of 7–21% divergence (maximum at 12%). Because of the greater access to highly divergent DNA, the increased potential contribution of mutator phenotypes to bacterial adaptation was substantially higher in a heterogeneous milieu (Figure 4) than in a homogeneous environment. In the model homogeneous environment, beneficial acquisitions were more likely overall, and mutators were still 300-fold more likely than nonmutators to acquire beneficial fragments. This contrast between the contribution of mutator and wild type to acquisition of beneficial divergent DNA was present even for the acquisition of a single amino acid change and became more marked as the number of amino acid substitutions required for beneficial effect was increased (Figure 5, A and B).

Mutator phenotypes, of course, also have an increased rate of mutation, including beneficial mutations, and this is an alternative adaptive route to a beneficial sequence change. Computer simulations, however, show that even rare genetic exchanges can accelerate evolution and undermine hypermutation as the dominant cause of adaptation (Tenaillon et al. 2000). Note also that the beneficial mutation must not be genetically linked to the mutator allele if it is to avoid accumulation of deleterious mutations (Taddei et al. 1997b; Funchain et al. 2000). Interestingly, Denamur et al. (2000) suggest that recombination may allow mutator phenotypes to restore defects in their DNA mismatch system and, hence, normalize their mutation and recombination frequencies; this might decrease the long-term cost of the mutator while still allowing for the short-term benefit in changing environments.

How does the probability of acquisition of a beneficial sequence by recombination compare to the probability of the same nucleotide changes occurring by simultaneous mutation within one generation? Of nucleotide point mutations at rate μ, about three-quarters result in changes of amino acids. Thus the rate of mutation of amino acids to a specific desired residue is 3μ/80 for q = 1/20 (1 over the number of amino acids possible). This is a generous estimate, because many amino
Figure 4.—The central 95% of probability density for acquisition of a beneficial DNA sequence in a homogeneous (solid line) and heterogeneous (shaded line) environment. Parameter values are $R_0 = 5 \times 10^{-5}$, $n = 3$, and $l = 300$. The probability densities as functions of nucleotide divergence are peaked at the nodules, but are nearly flat on this log 10 scale over the regions delimited. Note that a mutS$\Delta$ allele has $\omega = 16$, a wild type has $\omega = 64$, and a mutS$^{vp}$ strain, in which the wild-type mutS allele has been overexpressed, has $R_0 = 1.6 \times 10^{-5}$ and $\omega = 203$.

Acid changes cannot be accomplished by single-nucleotide mutations. But since this caveat holds for acquired differences as well, the comparison between the two remains germane. As an appropriate rate for gaining the beneficial sequence by simultaneous mutation alone, we therefore used $(3\mu/80)^n$. For our calculations, we used $\mu = 5.4 \times 10^{-10}/\text{bp}$ (Drake et al. 1998). The comparison of the probability of acquisition of a beneficial sequence by recombination to the probability of the same nucleotide changes occurring by simultaneous mutation revealed that protein sequence changes of 3–10 amino acids are many orders of magnitude more likely to result from recombination (Figure 6, A and B). We note that the large uncertainty in the overall rate of horizontal gene transfer, as parameterized by $R_0$, may have the effect of shifting these ratios downward by many orders of magnitude. However, the enormous ratios plotted in Figure 6 imply that this shift will alter only the crossover value of $n$, above which recombination will dominate, by at most one or two amino acids.

For both mutator and wild type, the dependence on $n$ and $\omega$ of the probability of beneficial acquisition is given approximately by

$$P_b \propto \left( \frac{n}{2eA_\omega X_0} \right)^n e^{-\omega/n/(4A_\omega X_0)^{1/4}}. \quad (9)$$

For $n \sim 5$, the dependence of this probability on the number of amino acid changes required primarily rests within the first factor, which is $\sim (1.6 \times 10^{-5})^n$. This may be compared with the much more rapidly decreasing dependence of the probability of arriving at the beneficial trait by simultaneous point mutations, $(3\mu/80)^n = (2.0 \times 10^{-10})^n$. In addition to diminished reliance on acquisitions involving small numbers of different amino acids for novel adaptation, the mutator gains probability of those acquisitions by a factor of almost $10^5$ over the wild-type bacteria. These comparisons depend only weakly on the parameters $A_\omega$ and $X_0$.

The element of time introduced by considering multiple generations of neutral mutation or recombination does not change the dynamics of adaptation, as both recombination and mutation events can occur sequentially in the same way. The relevant difference is in the ability of recombination to bring in multiple changes simultaneously. A set of mutations, for instance, may be acquired sequentially, if intermediate states are neutral.
**Figure 5.**—The log₁₀ probability of obtaining a beneficial phenotype by a given number, \( n \), of amino acid replacements, via a single beneficial recombination event in (A) a MutSΔ mutator strain and (B) a wild-type strain, in homogeneous (solid symbols) and heterogeneous (open symbols) environments. Open and solid circles correspond to a MutSΔ hypermutator \((R_0/\mu)\) and wild-type strain, respectively. Open and solid squares correspond to MutSΔ mutator strain and wild-type strain, respectively.

**Figure 6.**—The ratio of the probability of obtaining the beneficial phenotype by recombination vs. point mutation, (A) in the homogeneous environment (solid symbols) and (B) in the heterogeneous environment (open symbols). Open and solid circles correspond to a MutSΔ hypermutator \((R_0 = 5 \times 10^{-6}, \omega = 16, \mu = 5.4 \times 10^{-10})\). Open and solid squares correspond to wild type \((R_0 = 5 \times 10^{-6}, \omega = 64, \mu = 5.4 \times 10^{-10})\). Different basal rates of recombination \((R_0)\) proportionately shift the points as a whole up or down, but do not change the slopes delineated; thus for an organism with an \( R_0 \) of \( 5 \times 10^{-10} \), this graph would have all points lowered by a factor of 10⁴. \( R_0 \) has been measured in optimal laboratory conditions at a level \( 10^4 \) greater than that used for these plots (Vulic et al. 1997). The point mutation rate used was \( 5.4 \times 10^{-10} \) (Drake 1991).

However, in contrast to longer-term processes, such as sequential mutation, introduction of several nucleotides by a single recombination event allows the transit of valleys of dramatically reduced fitness as well as the transit of neutral ridges in genotype-fitness space.

An assumption, necessary for generality, inherent to the distributions of environmentally accessible DNA (Equation 2), is that the beneficial sequences are always present in the media, albeit at extremely low concentrations. Empirical studies addressing this distribution would be very valuable both for general models such as this and for more applied modeling of the uptake of exogenous DNA (Nielsen and Townsend 2001). Even when the beneficial sequence changes are not present together in any DNA sequence in the environment, it remains the case that multiple generations of recombina-
years (Ochman and Wilson 1987) has strong implications with regard to the evolution of antibiotic resistance in bacterial pathogens (Maynard Smith et al. 2000). We conclude that, given that beneficial sequences are present in the environment, mutators can play an important role in bacterial adaptation not only by providing mutations (Taddei et al. 1997a), but also by significantly expanding the sequence variability likely to be introduced into the bacterial genome. Thus, the environmental distribution and divergence of adaptive sequences available to bacteria are crucial to their capacity for rapid evolution by horizontal gene transfer.

We thank Dan Weinreich for comments on the manuscript. Support during this work was given to J.P.T. from the National Institutes of Health (NIH) and from a Harvard Merit Fellowship, to K.M.N. from The Research Council of Norway (Biodiversity Programme), to D.S.F. from the National Science Foundation, and to D.L.H. from the NIH.

LITERATURE CITED


Rayssiguier, C., D. S. Thaler and M. Radman, 1989 The barrier


Taddei, F., I. Matic, B. Godelle and M. Radman, 1997a To be a mutator, or how pathogenic and commensal bacteria can evolve rapidly. Trends Microbiol. 5: 427–428.


Communicating editor: M. Veuille

APPENDIX

We use several approximations based on small DNA divergence, $\nu$—the regime (dominant parameter range) for horizontal gene transfers—to derive analytic results regarding the divergence of beneficial acquisitions, as a function of $\omega$, $n$, and $l$. These results clarify the relative importance of factors such as gene length, $l$, number of substitutions needed for beneficial phenotype, $n$, and rate of recombination with divergence, parameterized by $\omega$, in affecting acquisition of beneficial DNA sequences.

From Equations 4 and 5, at small $\nu$,

$$a(\nu) = A_0 \nu^2 \quad \text{and} \quad x(\nu) = X_0 \nu^2 \quad \text{(A1)}$$

by discarding terms of higher order in $\nu$. For small $x$, the fact that $(1 - x)^b \approx e^{-bx}$ yields the following good approximations of Equations 2, 3, 7, and 8:

$$\epsilon(\nu) \approx C_0 \eta \ln v^{-\kappa \epsilon}$$

$$f(\nu, l) \approx 1 - e^{-A_0 X_0 l^{1/4}}$$

$$h(\nu) \approx A_0^{n} e^{2x \ln v^{-A_0^3 X_0^2 \sigma^2(l-v)}}$$

$$p_0(l, n, \nu) \approx C_0 R_0 (A_0 \eta)^n e^{-A_0 X_0 l^{1/4} v^{1-\omega_x - \kappa \nu - \eta \ln v + 2n \ln v}}.$$  \hspace{1cm} \text{(A2)}

When $n \ll l$, $(l - n)$ in Equations A2 may be approximated as $l$.

Setting aside the coefficient of the exponential in $p_0$, A2 is constant for any constant $n$, consider the term, $e^{-\epsilon}$, with

$$\epsilon = A_0 X_0 h^4 - \omega v - \kappa v + \eta \ln v + 2n \ln v,$$  \hspace{1cm} \text{(A3)}

which may be separated into three relevant parts

$$L = A_0 X_0 l,$$

$$\Omega = \omega + \kappa,$$

$$N = 2n + \eta.$$  \hspace{1cm} \text{(A4)}

Note that for most biological considerations, the term $L$ will vary most with $l$, $\Omega$ will vary with $\omega$, and $N$ will vary with $n$.

When $N$ is substantial, $p_0$ from Equations A2 will be quite sharply peaked as a function of $\nu$. For acquisitions of just a few amino acids, as modeled here, the peak is quite broad on a log scale (Figure 2). With large $L$, two regimes determine the peak location and width. If $\Omega$ is large, as it is for the wild type, $p_0(\nu)$ will be dominated by the $\Omega$ and $L$ terms, and the location of its peak, $\hat{\nu}$, will be at the minimum of $\epsilon$, where $\partial \epsilon / \partial \nu = \Omega - N/\nu = 0$, yielding

$$\hat{\nu} \approx \frac{N}{\Omega}.$$  \hspace{1cm} \text{(A5)}

If, on the other hand, $\Omega$ is small, as for a strong mutator, the peak will be at larger $\nu$, $p_0(\nu)$ will be dominated by the $N$ and $L$ terms, and $\hat{\nu}$ will be located where $\partial \epsilon / \partial \nu \approx 4L x^3 - N/\nu = 0$; thus

$$\hat{\nu} \approx \left( \frac{N}{4L} \right)^{1/4}.$$  \hspace{1cm} \text{(A6)}

A weaker mutator resides between these two regimes.

As the number of amino acids to be acquired increases, both the mutator and the wild type are in the regime of relatively small $\Omega$. The dominant $n$ and $\omega$ dependence of the probability of beneficial acquisition can then be obtained by integrating $p_0$ in a Gaussian approximation around its peak at $\hat{\nu}$. This yields

$$\int \sqrt{\frac{N}{4\pi L}} e^{-\Omega(N/L)^{1/4}},$$

as reported in the text.