

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 DYKHUIZEN 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; DOWSON *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 DYKHUIZEN 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*; RAYSIGUIER *et 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 gene 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-

¹Corresponding author: Department of Plant and Microbial Biology, 321 Koshland Hall, University of California, Berkeley, CA 94720. E-mail: townsend@nature.berkeley.edu

²Present address: Department of Pharmacy, School of Medicine, University of Tromsø, N-9037 Tromsø, Norway.

TABLE 1
Definitions of parameters used in the model

Parameter	Range of study	Definition
v	0–1	Nucleotide divergence
ω	16–203	Exponential rate of decrease of recombination with sequence divergence
l	0–1000	Length of coding DNA fragment, in codons
n	1–10	Number of amino acid changes required for a beneficial effect
q	0.05–0.067	Probability that a differing amino acid is the desired residue
η	0–3	First parameter for the beta distribution $c(v)$
κ	1–9	Second parameter for the beta distribution $c(v)$
R_0	1.6×10^{-7} – 5×10^{-6}	Rate of recombination at zero divergence
C_0	10–20	Normalization constant for $c(v)$, equal to $\Gamma(\eta + \kappa + 2)/(\Gamma(\eta + 1) \times \Gamma(\kappa + 1))$
A_0	0–4	Parameter for $x(v)$ describing probability that diverged amino acids are deleterious
X_0	0–4	Parameter for $a(v)$ describing amino acid divergence

bine with significantly divergent (>20%) chromosomal DNA as compared to wild-type phenotypes, which recombine primarily with DNA of low divergence (\sim 1–2%). Indeed, mutator phenotypes are amplified to high levels after selection for interspecies recombination events in laboratory populations of *Salmonella* sp. and *E. coli* (FUNCHAIN *et al.* 2001).

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

$$r(v) = R_0 e^{-\omega v}, \quad (1)$$

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

The link between recombinogenic mutators with small values of ω (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 hypothesize that rare recombination events facilitated by recombinogenic mutator bacteria provide a source of variation virtually inaccessible by point muta-

tions 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; DRÖGE *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 β -function

$$c(v) = C_0 v^\eta (1 - v)^\kappa, \quad (2)$$

where C_0 is a normalization constant, and η and κ 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 2
Definitions of functions derived in the model

Function	Definition
$r(v)$	Rate of acquisition of DNA sequence by recombination
$c(v)dv$	Fraction of ambient DNA in interval $(v, v + dv)$
$f(v, l)$	Probability that an acquired fragment has a deleterious effect
$h(v, l, n)$	Probability that an acquired fragment has a beneficial effect
$\rho_a(v, l)dv$	Probability of a deleterious recombination event, within interval $(v, v + dv)$
$\rho_b(v, l)dv$	Probability of a beneficial recombination event, within interval $(v, v + dv)$

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,

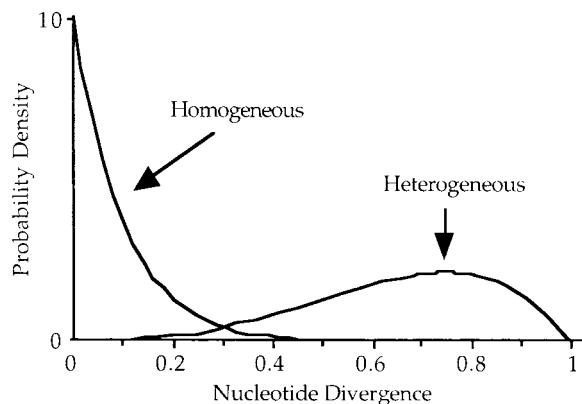


FIGURE 1.—Probability density, $c(v)$, of accessible DNA as a function of sequence divergence in a model homogeneous or heterogeneous environment. The total concentrations of accessible DNA were equal in the two environments; the curves both integrate to unity.

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

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 $x(v) \propto a(v)$ 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 :

$$x(v) = X_0v^2(1 - v) + v^4. \quad (5)$$

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

$x(v)$ that are of order v^2 for small v would not change the dynamics described here.

The probability of a deleterious recombination event occurring in the DNA environment specified by $c(v)$ is

$$\rho_d(v, l) = f(v, l)r(v)c(v). \quad (6)$$

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

increasing sequence divergence. The size of fragment considered here, ~ 900 bp, is biologically relevant.

To estimate the probability that a beneficial combination of nucleotides was present in acquired DNA of divergence v , we assumed that n specific amino acid sites must be replaced by n other specific amino acids in a particular gene and that this acquisition had to be accompanied by no deleterious replacements in the neighboring $l - n$ sites. The probability that such a replacement has the prescribed beneficial properties is

$$h(v, l, n) = q^n (a(v))^n (1 - a(v))^n (1 - a(v)x(v))^{l-n}, \quad (7)$$

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 $q \sim 1/15$ (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

$$\rho_b(v, l, n) = h(v, l, n)r(v)c(v). \quad (8)$$

The total probability of a beneficial recombination event in relation to the distribution of accessible sequences is calculated as $\int_0^1 \rho_b(v, l, n) dv$, and that of a deleterious recombination event as $\int_0^1 \rho_d(v, l) dv$. An analysis based on small v revealed the degree to which these results depend on the parameters ω , 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

$$\frac{2n + \eta}{\omega + \kappa},$$

whereas for a very strong mutator, the peak is at

$$\left(\frac{2n + \eta}{A_0 X_0 l}\right)^{1/4}.$$

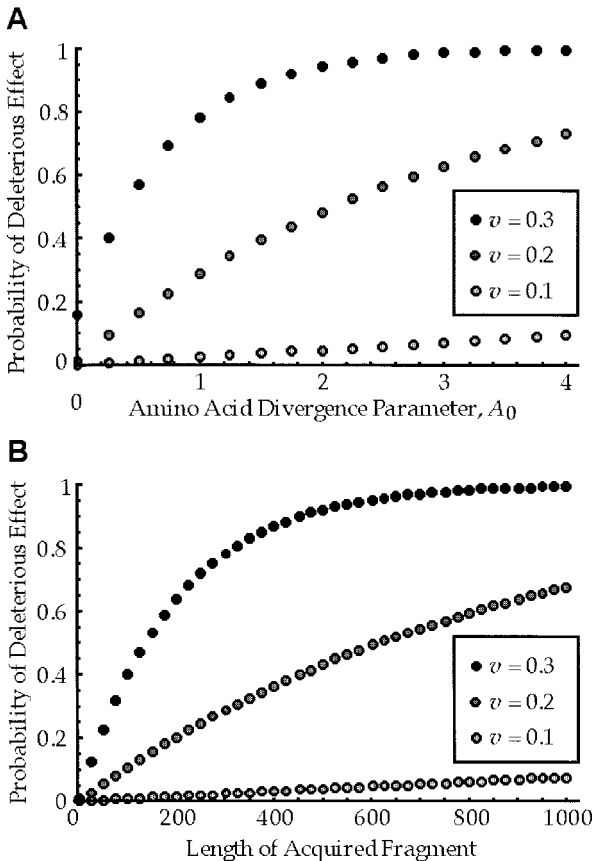


FIGURE 2.—Modeled probability that an integrated DNA fragment is deleterious if it has a nucleotide divergence of $v = 0.1$ (solid), 0.2 (dark shading), or 0.3 (light shading), (A) for increasing numbers of amino acid changes per nucleotide change ($0 \leq A_0 < 4$), and (B) for increasing lengths ($0 \leq l \leq 1000$) of the recombining DNA fragment.

RESULTS AND DISCUSSION

Because recombination increases both deleterious and beneficial sequence acquisitions, we have focused on the adaptive importance of single events rather than

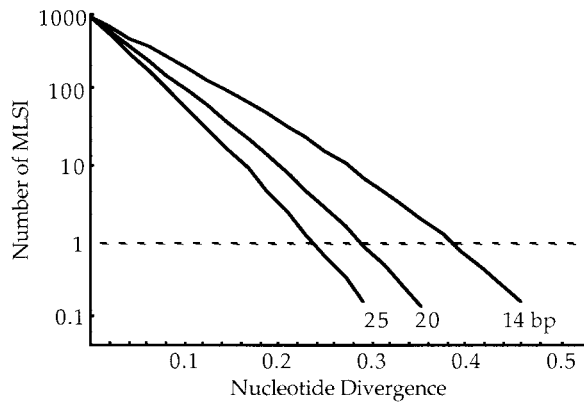


FIGURE 3.—Expected number of minimum lengths of sequence identity (MLSI), with 14-, 20-, or 25-bp segments of absolute sequence identity, in a 900-bp sequence. The dashed line indicates an expectation of one MLSI.

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 ($R_0 = 0.05$, $\omega = 64$) and mutator *mutS* Δ ($R_0 = 0.05$, $\omega = 16$). 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 ($R_0 = 0.0016$, $\omega = 203$). 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 $\sim 30-35\%$ 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 $(3l - m + 1)(1 - v)^m$, 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^{-mv} , suggesting that the minimum value of ω is determined by the length, m , of a MLSI. The observed value of $\omega = 16$ for a *mutS* Δ 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 $\rho_b(v)$ of acquisition of $n = 3$ specific amino acids in a fragment of length $l = 300$ amino acids from a donor with divergence v (ZAWADSKI 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 $3\mu/80$ for $q = 1/20$ (1 over the number of amino acids possible). This is a generous estimate, because many amino

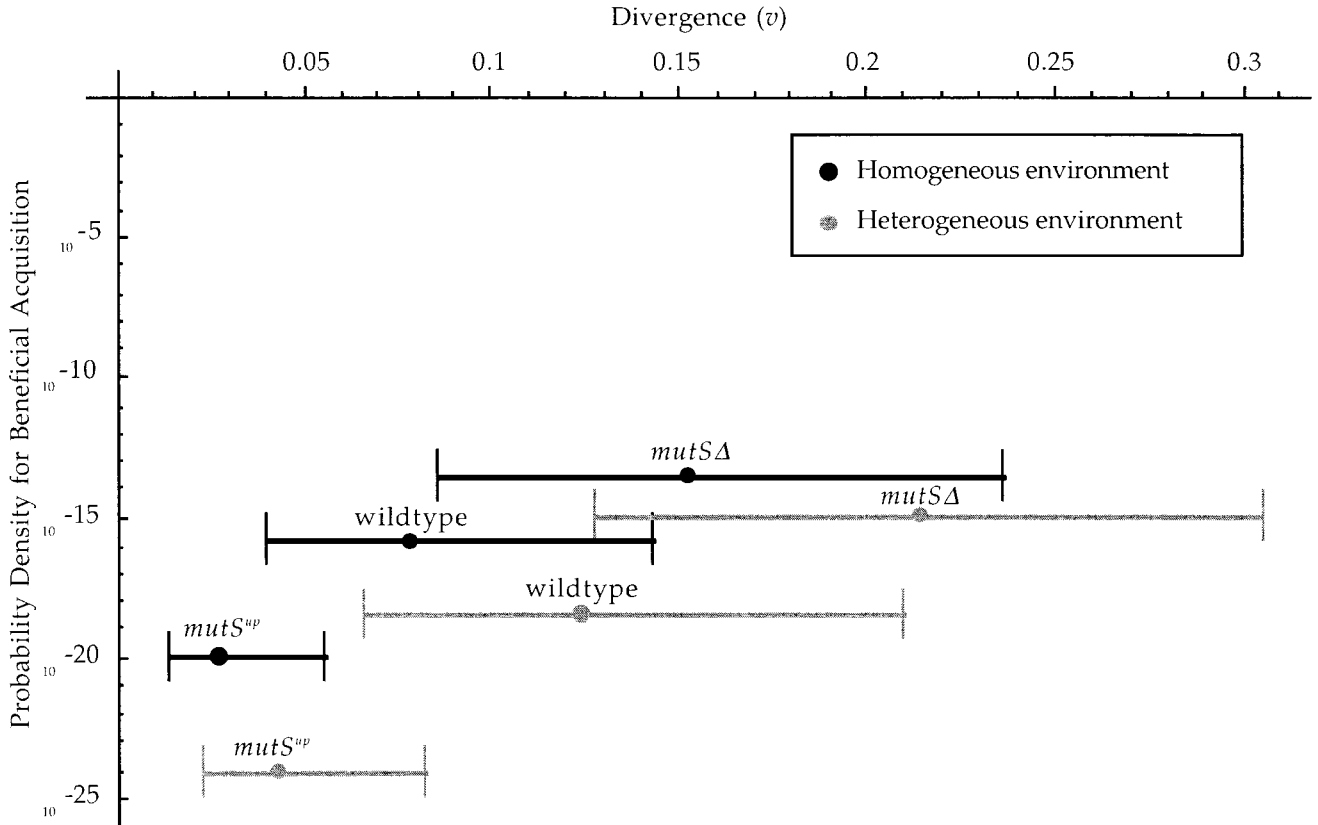


FIGURE 4.—The central 95% of probability density for acquisition of a beneficial DNA sequence in a homogeneous (solid line) and heterogeneous (shaded line) environment. Parameter values are $R_0 = 5 \times 10^{-6}$, $n = 3$, and $l = 300$. The probability densities as functions of nucleotide divergence are peaked at the nodules, but are nearly flat on this \log_{10} scale over the regions delimited. Note that a *mutSA* allele has $\omega = 16$, a wild type has $\omega = 64$, and a *mutS^{up}* strain, in which the wild-type *mutS* allele has been overexpressed, has $R_0 = 1.6 \times 10^{-7}$ 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 $\mu = 5.4 \times 10^{-10}$ /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 ω of the probability of beneficial acquisition is given approximately by

$$P_b \propto \left(q \sqrt{\frac{n}{2eA_0X_0l}} \right)^n e^{-\omega(n/4A_0X_0l)^{1/4}}. \quad (9)$$

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 $\sim(1.6 \times 10^{-3})^n$. This may be compared with the much more rapidly decreasing dependence of the probability of arriving at the beneficial trait by simultaneous point mutations, $(3\mu/80)^n \approx (2.0 \times 10^{-11})^n$. 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_0 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.

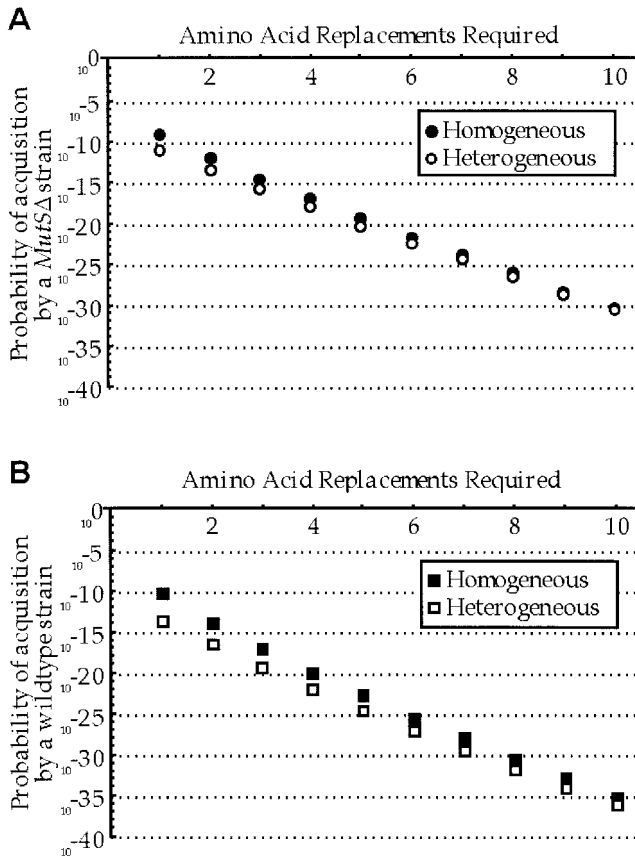


FIGURE 5.—The \log_{10} probability of obtaining a beneficial phenotype conferred by a given number, n , of amino acid replacements, via a single beneficial recombination event in (A) a *MutSA* mutator strain and (B) a wild-type strain, in homogeneous (solid symbols) and heterogeneous (open symbols) environments. Open and solid circles correspond to a *mutSA* hypermutator ($R_0 = 5 \times 10^{-6}$, $\omega = 16$, $\mu = 5.4 \times 10^{-8}$). Open and solid squares correspond to wild type ($R_0 = 5 \times 10^{-6}$, $\omega = 64$, $\mu = 5.4 \times 10^{-10}$). Note that probabilities on the order of 10^{-20} are infrequent but possible, whereas probabilities below 10^{-40} are biologically unrealistic for any population, even over geological time scales.

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 recombi-

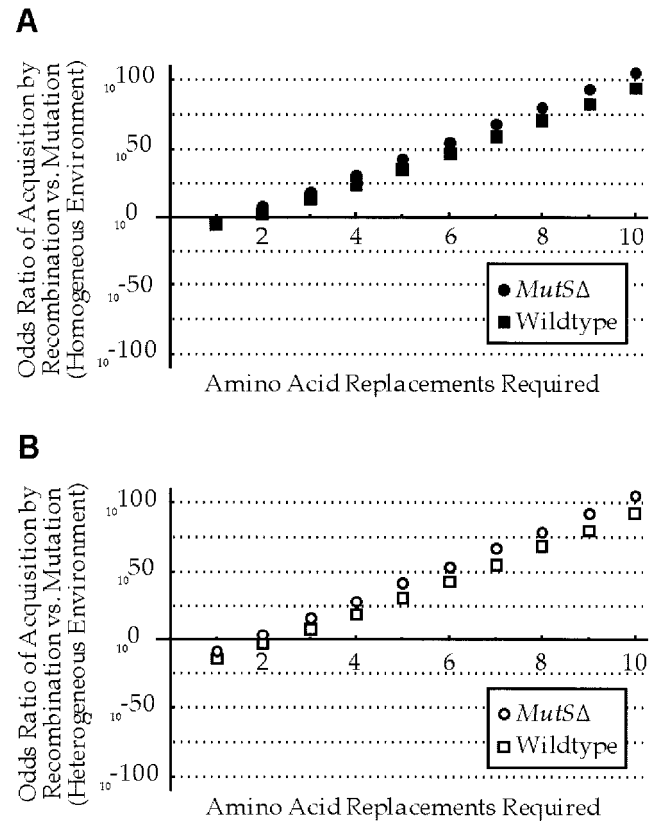


FIGURE 6.—The ratio of the probability of obtaining the beneficial phenotype by recombination *vs.* point mutation, (A) in the homogeneous environment (solid symbols) and (B) in the heterogeneous environment (open symbols). Open and solid circles correspond to a *mutSA* hypermutator ($R_0 = 5 \times 10^{-6}$, $\omega = 16$, $\mu = 5.4 \times 10^{-8}$). 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 5×10^{-11} , this graph would have all points lowered by a factor of 10^5 . 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 5.4×10^{-10} (DRAKE 1991).

nation may lead to the creation of the postulated adaptive sequence. It is clear that the existing diversity in the ecosystem can be recombined successively to produce highly adapted proteins (STEMMER 1994). Furthermore, the genetic diversity present in the environment is presumably more likely to be beneficial than would random DNA or DNA mutated at random. Further population genetic modeling of the potential for sequential uptake events with an adaptive trajectory of specified selection coefficients is needed. The explanation of the observed high frequencies of mutator alleles in selected natural environments via this kind of model deserves further attention.

That bacteria can effectively acquire DNA sequences that have been separately evolving for up to 150 million

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

- BESSEN, D. E., and S. K. HOLLINGSHEAD, 1994 Allelic polymorphism of *emm* loci provides evidence for horizontal gene spread in group A streptococci. *Proc. Natl. Acad. Sci. USA* **91**: 3280–3284.
- BLATTNER, F. R., G. PLUNKETT, III, C. A. BLOCK, N. T. PERNA, V. BURLAND *et al.*, 1997 The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453–1462.
- BOWLER, L. D., Q.-Y. ZHANG, J.-Y. RIOU and B. G. SPRATT, 1994 Interspecies recombination between the *penA* genes of *Neisseria meningitidis* and commensal *Neisseria* species during the emergence of penicillin resistance in *N. meningitidis*: natural events and laboratory simulation. *J. Bacteriol.* **176**: 333–337.
- BROWN, E. W., J. E. LECLEERC, B. LI, W. L. PAYNE and T. A. CEBULA, 2001 Phylogenetic evidence for horizontal transfer of *mutS* alleles among naturally occurring *Escherichia coli* strains. *J. Bacteriol.* **183**: 1631–1634.
- CLAVERYS, J.-P., M. PRUDHOMME, I. MORTIER-BARRIÈRE and B. MARTIN, 2000 Adaptation to the environment: *Streptococcus pneumoniae*, a paradigm for recombination-mediated genetic plasticity? *Mol. Microbiol.* **35**: 251–259.
- COYNE, J. A., and H. A. ORR, 1989 Patterns of speciation in *Drosophila*. *Evolution* **43**: 362–381.
- DAVISON, J., 1999 Genetic exchange between bacteria in the environment. *Plasmid* **42**: 73–91.
- DENAMUR, E., G. LECOINTRE, P. DARLU, O. TENAILLON, C. ACQUAVIVA *et al.*, 2000 Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. *Cell* **103**: 711–721.
- DOWSON, C. G., A. HUTCHISON, J. A. BRANNIGAN, R. C. GEORGE, D. HANSMAN *et al.*, 1989 Horizontal transfer of penicillin-binding protein genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. USA* **86**: 8842–8846.
- DOWSON, C. G., A. HUTCHISON, N. WOODFORD, A. P. JOHNSON, R. C. GEORGE *et al.*, 1990 Penicillin-resistant viridans streptococci have obtained altered penicillin-binding protein genes from penicillin-resistant strains of *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. USA* **87**: 5858–5862.
- DRAKE, J. W., 1991 A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl. Acad. Sci. USA* **88**: 7160–7164.
- DRAKE, J. W., B. CHARLESWORTH, D. CHARLESWORTH and J. F. CROW, 1998 Rates of spontaneous mutation. *Genetics* **148**: 1667–1686.
- DRÖGE, M., A. PUHLER and W. SELBITSCHKA, 1999 Horizontal gene transfer among bacteria in terrestrial and aquatic habitats as assessed by microcosms and field studies. *Biol. Fertil. Soils* **29**: 221–245.
- FEIL, E., M. C. J. MAIDEN, M. ACHTMAN and B. G. SPRATT, 1999 The relative contributions of recombination and mutation to the divergence of clones of *Neisseria meningitidis*. *Mol. Biol. Evol.* **16**: 1496–1502.
- FEIL, E. J., J. MAYNARD SMITH, M. C. ENRIGHT and B. G. SPRATT, 2000 Estimating recombinational parameters in *Streptococcus pneumoniae* from multilocus sequence typing data. *Genetics* **154**: 1439–1450.
- FENG, D.-F., G. CHO and R. F. DOOLITTLE, 1997 Determining divergence times with a protein clock: update and reevaluation. *Proc. Natl. Acad. Sci. USA* **94**: 13028–13033.
- FUNCHAIN, P., A. YEUNG, J. L. STEWART, R. LIN, M. M. SLUPSKA *et al.*, 2000 The consequences of growth of a mutator strain of *Escherichia coli* as measured by loss of function among multiple gene targets and loss of fitness. *Genetics* **154**: 959–970.
- FUNCHAIN, P., A. YEUNG, J. STEWART, W. M. CLENDENIN and J. H. MILLER, 2001 Amplification of mutator cells in a population as a result of horizontal transfer. *J. Bacteriol.* **183**: 3737–3741.
- GUTTMAN, D. S., and D. E. DYKHUIZEN, 1994 Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science* **266**: 1380–1383.
- JAIN, R., M. C. RIVERA and J. A. LAKE, 1999 Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl. Acad. Sci. USA* **96**: 3801–3806.
- KROLL, J. S., K. E. WILKS, J. L. FARRANT and P. R. LANGFORD, 1998 Natural genetic exchange between *Haemophilus* and *Neisseria*: intergeneric transfer of chromosomal genes between major human pathogens. *Proc. Natl. Acad. Sci. USA* **95**: 12381–12385.
- 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.
- MAJEWSKI, J., and F. M. COHAN, 1998 The effect of mismatch repair and heteroduplex formation on sexual isolation in *Bacillus*. *Genetics* **148**: 13–18.
- MAJEWSKI, J., and F. COHAN, 1999 DNA sequence similarity requirements for interspecific recombination in *Bacillus*. *Genetics* **153**: 1525–1533.
- MAJEWSKI, J., P. ZAWADZKI, P. PICKERILL, F. M. COHAN and C. G. DOWSON, 2000 Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. *J. Bacteriol.* **182**: 1016–1023.
- 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., 1995 Interspecies gene exchange in bacteria: the role of SOS and mismatch repair systems in evolution of species. *Cell* **80**: 507–515.
- MATIC, I., F. TADDEI and M. RADMAN, 1996 Genetic barriers among bacteria. *Trends Microbiol.* **4**: 69–73.
- 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., E. J. FEIL and N. H. SMITH, 2000 Population structure and evolutionary dynamics of pathogenic bacteria. *BioEssays* **22**: 1115–1122.
- NIELSEN, K. M., and J. P. TOWNSEND, 2001 Environmental exposure, horizontal transfer, and selection of transgenes in bacterial populations, pp. 145–158 in *Enhancing Biocontrol Agents and Handling Risks*, edited by M. VURRO, J. GRESSEL, T. BUTT, G. E. HARMAN, A. PILGERAM *et al.* IOS Press, Amsterdam.
- NIELSEN, K. M., M. VAN WEERELT, T. N. BERG, A. M. BONES, A. N. HAGLER *et al.*, 1997 Natural transformation and availability of transforming chromosomal DNA to *Acinetobacter calcoaceticus* in soil microcosms. *Appl. Environ. Microbiol.* **63**: 1945–1952.
- NIELSEN, K. M., K. SMALLA and J. D. VAN ELSAS, 2000 Natural transformation of *Acinetobacter* sp. strain BD413 with cell lysates of *Acinetobacter* sp., *Pseudomonas fluorescens* and *Burkholderia cepacia* in soil microcosms. *Appl. Environ. Microbiol.* **66**: 206–212.
- NOTLEY-MCROBB, L., S. SEETO and T. FERENCI, 2002 Enrichment and elimination of *mutY* mutators in *Escherichia coli* populations. *Genetics* **162**: 1055–1062.
- OCHMAN, H., and A. C. WILSON, 1987 Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.* **26**: 74–86.
- OCHMAN, H., J. G. LAWRENCE and E. A. GROISMAN, 2000 Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299–304.
- OLIVER, A., R. CANTÓN, P. CAMPO, F. BAQUERO and J. BLÁZQUEZ, 2000 High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science* **288**: 1251–1253.
- ORR, H. A., 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* **139**: 1805–1813.
- RAYSSIGUIER, C., D. S. THALER and M. RADMAN, 1989 The barrier

to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* **342**: 396–401.

- RIVERA, M. C., R. JAIN, J. E. MOORE and J. A. LAKE, 1998 Genomic evidence for two functionally distinct gene classes. *Proc. Natl. Acad. Sci. USA* **95**: 6239–6244.
- ROBERTS, M. S., and F. M. COHAN, 1995 Recombination and migration rates in natural populations of *Bacillus subtilis* and *Bacillus mojavensis*. *Evolution* **79**: 1081–1094.
- SHEN, P., and H. V. HUANG, 1986 Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics* **112**: 441–457.
- SPRATT, B. G., 1994 Resistance to antibiotics mediated by target alterations. *Science* **264**: 388–393.
- SPRATT, B. G., Q.-Y. ZHANG, D. M. JONES, A. HUTCHISON, J. A. BRANNIGAN *et al.*, 1989 Recruitment of a penicillin-binding protein gene from *Neisseria flavescens* during the emergence of penicillin resistance in *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. USA* **86**: 1989.
- STEMMER, W. P. C., 1994 DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91**: 10747–10751.
- 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.
- TADDEI, F., M. RADMAN, J. MAYNARD-SMITH, B. TOUPANCE, P. H. GOUYON *et al.*, 1997b Role of mutator alleles in adaptive evolution. *Nature* **387**: 700–702.
- 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.
- VULIC, M., F. DIONISIO, F. TADDEI and M. RADMAN, 1997 Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. *Proc. Natl. Acad. Sci. USA* **94**: 9763–9767.
- WATT, V. M., C. J. INGLES, M. S. URDEA and W. J. RUTTER, 1985 Homology requirements for recombination in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **82**: 4768–4772.
- ZAWADSKI, P., and F. M. COHAN, 1995 The size and continuity of DNA segments integrated in *Bacillus* transformation. *Genetics* **141**: 1231–1243.
- ZAWADSKI, P., M. S. ROBERTS and F. M. COHAN, 1995 The log-linear relationship between sexual isolation and sequence divergence in *Bacillus* transformation is robust. *Genetics* **140**: 917–932.

Communicating editor: M. VEUILLE

APPENDIX

We use several approximations based on small DNA divergence, v —the regime (dominant parameter range) for horizontal gene transfers—to derive analytic results regarding the divergence of beneficial acquisitions, as a function of ω , 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 ω , in affecting acquisition of beneficial DNA sequences.

From Equations 4 and 5, at small v ,

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

by discarding terms of higher order in v . For small x ,

the fact that $(1 - x)^k \approx e^{-kx}$ yields the following good approximations of Equations 2, 3, 7, and 8:

$$c(v) \approx C_0 e^{\eta \ln v - \kappa v}$$

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

$$h(v) \approx A_0^n e^{2n \ln v - A_0 v^2 X_0 v^{2(l-n)}},$$

$$\rho_b(l, n, v) \approx C_0 R_0 (A_0 q)^n e^{-A_0 X_0 (l-n)v^{4l} - \omega v - \kappa v + \eta \ln v + 2n \ln v}. \quad (\text{A2})$$

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

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

$$\varepsilon \equiv A_0 X_0 l v^4 - \omega v - \kappa v + \eta \ln v + 2n \ln v, \quad (\text{A3})$$

which may be separated into three relevant parts

$$L \equiv A_0 X_0 l,$$

$$\Omega \equiv \omega + \kappa,$$

$$N \equiv 2n + \eta. \quad (\text{A4})$$

Note that for most biological considerations, the term L will vary most with l , Ω will vary with ω , and N will vary with n .

When N is substantial, ρ_b from Equations A2 will be quite sharply peaked as a function of v . 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 Ω is large, as it is for the wild type, $\rho_b(v)$ will be dominated by the Ω and L terms, and the location of its peak, \hat{v} , will be at the minimum of ε , where $\partial\varepsilon/\partial v \approx \Omega - N/v = 0$, yielding

$$\hat{v} \approx \frac{N}{\Omega}. \quad (\text{A5})$$

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

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

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 Ω . The dominant n and ω dependence of the probability of beneficial acquisition can then be obtained by integrating ρ_b in a Gaussian approximation around its peak at \hat{v} . This yields

$$\left(q\sqrt{\frac{N}{4eL}}\right)^{N/2} e^{-\Omega(N/4L)^{1/4}}$$

as reported in the text.

