Transiently Beneficial Insertions Could Maintain Mobile DNA Sequences in Variable Environments

Richard J. Edwards† and John F. Y. Brookfield**

*Institute of Genetics, University of Nottingham, Queens Medical Centre, Nottingham, United Kingdom; and †Clinical Pharmacology, Royal College of Surgeons in Ireland, Dublin, Ireland

The maintenance of mobile DNA sequences in clonal organisms has been seen as a paradox. If selfish mobile sequences spread through genomes only by overreplication in transposition, then sexuality is necessary for their spread through populations. The persistence of bacterial transposable elements without obvious dominant selectable markers has previously been explained by horizontal transfer. However, advantageous insertions of mobile DNAs are known in bacteria. Here we model maintenance of an otherwise selfish mobile DNA element in a clonal species in which selection for null mutations occurs during one of two temporally alternating environments. Large areas of parameter space permit maintenance of mobile DNAs where, without selection, they would have gone extinct. Horizontal transfer diminishes, rather than enhances, mean copy number. In finite populations, effective population sizes are greatly reduced by selective sweeps, and mean copy number can be increased as the reduced variance in copy number results in reduced selection.

Introduction

There is an intimate connection between sexual reproduction and the maintenance of mobile DNA sequences. If mobile DNAs are invariably harmful, then they will lower the fitness of their hosts. In a clonal population that contains some cells which lack transposable elements altogether, this genotype will inexorably spread, replacing all transposable-element–containing cells, however much the transposable element copy number increases in these cells (Cavalier-Smith 1980; Hickey 1982). Instead, the maintenance of selfish mobile DNAs in clonal organisms requires some horizontal spread of the mobile DNAs. The low rate of horizontal transfer seen in most bacteria is, however, probably insufficient to counter the loss of elements by selection against their harmful effects (Condit, Stewart, and Levin 1988). For these reasons, we must suspect that the maintenance of mobile DNAs in such primarily clonal organisms as the eubacteria is the result of selection. In the case of composite transposons this selection is for mobilization of the exogenous genes that these elements contain. Bacterial insertion sequences, however, have no dominant selectable function, and selection to maintain these sequences must be of a different kind. Here we model a novel mechanism for the maintenance of such mobile sequences in a clonal population through their capacity to generate insertion mutations and to interrupt gene function transiently in a fluctuating environment.

The potential for successful mutator strains in bacteria is well documented (Chao and Cox 1983; Mao et al. 1997; Sniegowski, Gerrish, and Lenski 1997; Taddei et al. 1997; Tenaillon et al. 1999), and bacterial transposable elements can sometimes act as beneficial mutators (Chao et al. 1983; Chao and McBroom 1985; Modi et al. 1992; Blot 1994). In some situations, insertions of insertion sequences into functional genes create advantageous genotypes in laboratory environments (Zinser and Kolter 2000). We postulate that corresponding changes will be possible in the wild, given the great range of environments that bacterial populations encounter. However, if advantageous insertion mutations into functional genes maintain insertion sequences, selection for such insertions will, with time, inactivate all such genes whose loss is advantageous. Thus, one would expect that, in organisms selected for efficiency such as bacteria, these inactivated genes will be lost, the process will stop, and, ultimately, insertion sequences themselves will be lost. The only viable model for the maintenance of mobile DNAs through their ability to create advantageous null mutations requires the selective advantages of insertions to be environment-dependent, such that in a different, future environment there would be selection for reactivation of the gene. It follows, then, that these mobile DNA insertions must be capable of reversion by precise excision events. Such excisions are seen in nature (Kleckner 1981; Allgood and Silhavy 1988; Mills et al. 1992; Hammerschidt et al. 1996; Ziebuhr et al. 1999), and there is even some evidence that these excisions may be induced by environmental stress (Aleshkin, Kadzhaev, and Markov 1998). Many authors have identified reversible insertions of insertion sequences that are of benefit to adapting bacteria. Examples include colonial variation of *Shigella flexneri* (Mills et al. 1992), modulation of cell surface sialic acid expression in *Neisseria meningitidis* (Hammerschidt et al. 1996), and phase variation of biofilm formation in *Staphylococcus epidermidis* (Ziebuhr et al. 1999).

Key words: transient insertion mutations, transposable elements, mobile DNA, insertion sequences.

E-mail: john.brookfield@nottingham.ac.uk.

Mol. Biol. Evol. 20(1):30–37. 2003
DOI: 10.1093/molbev/msg001

Materials and Methods

Basic Model

The population is defined by the frequencies of classes of cells, $p[I, J]$, where I and J are the number of transposable elements (TEs) at the site of interest and neutral sites, respectively. The population experiences T generations in each of two alternating environments e, which respectively favor presence ($e = 1$) or absence ($e = 2$) of insertions at the site of interest, with frequency changes for classes being calculated each generation. This continues until equilibrium is reached, where class frequencies are the same after consecutive cycles, or until

the TE is lost from the population. Note that the term *generation* does not necessarily have to correspond to a single cell division cycle. Instead, it is merely an arbitrary length of time during which the appropriate frequency changes occur. Because transposition, deletion, and cell death can be independent of chromosomal DNA replication, these generations could occur in exponential growth or stationary phases.

Each generation the following events occur:

(A) *Deletions*. Each class $[I, J]$ experiences deletion (precise excision) of elements at a rate of $\mu(I+J)$, and make contributions to new class frequencies as follows:

(i) $p[I, J](1 - \mu(I + J)) \Rightarrow p[I, J]$;
(ii) $p[I, J]\mu\, I \Rightarrow p[I-1, J]$;
(iii) $p[I, J]\mu\, J \Rightarrow p[I, J-1]$.

(B) *Fitness*. Each class $[I, J]$ experiences selection of t per element and, depending on I and e, selection s at the site of interest:

(i) When $I = 0$ and $e = 1$, $p[I, J](1 - t)^J(1 - s)$;
(ii) When $I > 0$ and $e = 2$, $p[I, J](1 - t)^{(I+J)}(1 - s)$;
(iii) All other cases, $p[I, J](1 - t)^{(I+J)}$.

(C) *Transposition*. Each class $[I, J]$ experiences transposition of elements at a rate of $\upsilon(I + J)$. Transposition targets are the site of interest, i, neutral sites, n, or deleterious sites, $(1 - i - n)$. Classes make contributions to new class frequencies as follows:

(i) $p[I, J](1 - \upsilon(I + J)) \Rightarrow p[I, J]$;
(ii) $p[I, J]\upsilon(I + J)i \Rightarrow p[I + 1, J]$;
(iii) $p[I, J]\upsilon(I + J)n \Rightarrow p[I, J + 1]$.

The remaining $p[I, J]\upsilon(I + J)(1 - i - n)$ are lost through deleterious insertions and effectively add to the fitness cost, t.

(D) *Normalization*. Classes with a frequency under 10^{-20} are removed. The frequencies of the remaining classes are then divided by the sum of all classes to restore a total frequency of 1: $p[I, J]/\Sigma p[I, J]$.

Horizontal Transmission

For horizontal transmission, a certain proportion of transpositions h generate insertions in random members of the population, rather than the donor cell. In step (C), contributions (ii) and (iii) are altered by a factor $(1 - h)$ giving a probability of receiving a horizontal transmitted element, $H = \Sigma p[I, J]\upsilon(I + J)h(i + n)$. Following an additional normalization step, $p[I, J]/\Sigma p[I, J]$ and $H/\Sigma p[I, J]$, this "pool" of insertions is then distributed among the population. Because lethal insertions will affect every class equally, these are ignored and H adjusted accordingly, $H / (1 - H(1 - i - n))$. Contributions to new class frequencies are then made:

(i) $p[I, J](1 - H) \Rightarrow p[I, J]$;
(ii) $p[I, J]Hi \Rightarrow p[I + 1, J]$;
(iii) $p[I, J]Hn \Rightarrow p[I, J + 1]$.

The final normalization and class removal (step D) then occurs as before.

Finite Population Size Models

For finite population models, class frequencies are converted into numbers of cells during step (D). An expected number $N_E[I, J]$ is first determined by multiplying frequencies by the population size N, $N_E[I, J] = p[I, J]N$. These are then converted to actual numbers $N_A[I, J]$. For $N_E[I, J] > 100$, $N_A[I, J] = N_E[I, J]$. For $N_E[I, J] \leqslant 100$, the model incorporates random genetic drift, where $N_A[I, J]$ is Poisson distributed with a mean given by $N_E[I, J]$. The population is then returned to frequencies: $p[I, J] = N_A[I, J]/\Sigma N_A[I, J]$.

Analytical Approximation of Minimum Cycle Length

The cycle length, T, allowing stability falls within a range determined by the parameters, s, n, i, t, υ, and μ. A simplified model can be considered in which there is never more than one element at the site of interest and never more than one element at a neutral site. We call an individual with no elements (0,0); one with an element only at the neutral site, (0,1); one with an element only at the site of interest, (1,0); and one at both sites, (1,1).

If the cycle time is short the element will not be eliminated from the site of interest during environment 2, nor will it spread to fixation during environment 1. The element at the site of interest is subject to selection of strength $s - t$ during environment 1 and $-s - t$ during environment 2, which lasts for an equal length. Without transposition or deletion the element would inexorably be lost at the site of interest. However, transposition into the site of interest may allow the element to persist stably. If we assume, for now, that throughout the cycle there is exactly one element present at a neutral site (i.e., all cells are of the (0,1) or (1,1) types) it is possible to model the cycling of the frequencies of these (0,1) and (1,1) cells. Cells with an insertion at the site of interest have frequency p and a selective advantage in environment 1, which we call α. Cells of the (0,1) class receive transpositions into the site of interest at rate $i\upsilon$. We call β the selective disadvantage of the element in the site of interest in environment 2.

Integration of frequency changes in the two environments, and setting p at the end of environment 2 to be the same as at the start of environment 1 (defining an equilibrium cycle for p) requires p to be $i\upsilon(\alpha e^{\beta T} + \beta e^{\alpha T})/(\alpha\beta(e^{\beta T} - e^{\alpha T}))$ at the start of environment 1 and p to be given by

$$p/(1-p) = i\nu e^{\alpha T}(\alpha e^{\beta T} + \beta e^{\alpha T})/(\alpha^2\beta(e^{\beta T} - e^{\alpha T})) \quad (1)$$

at the start of environment 2. The transposition rate during the cycle assumes that there is one element at a neutral site. Thus, all cells without the element at the site of interest (which will be almost all cells at the start of environment 1) are assumed to be (0,1) cells and not (0,0) cells. Because selection at rate $t + (1 - n - i)\upsilon$ will constantly favor (0,0) cells relative to (0,1) cells, maintaining a high proportion of (0,1) cells requires input from the (1,1) class, arising by deletion of the element at the site of interest.

As an approximation, we assume that, at the start of environment 1, one half of the cells lacking an element at

the site of interest are (0,1) and the other half are (0,0). Selection against the former cells will mean that at the start of environment 2 a proportion $1/(1 + e^{(t+(1-n-i)\nu)T})$ of the cells lacking the element at the site of interest will be (0,1). Thus for an equilibrium cycle the frequency of (0,1) cells among the (0,0) and (0,1) cells has to be raised, at the time when the proportion of cells at the site of interest is high (at the end of environment 1 and the start of environment 2), from $1/(1 + e^{(t+(1-n-i)\upsilon)T})$ to $e^{(t+(1-n-i)\upsilon)T}/(1 + e^{(t+(1-n-i)\upsilon)T})$. This happens by deletion from the site of interest, which converts (1,1) cells, which have frequency p, to (0,1) cells. If $p \gg (1 - p)$ at the end of environment 1, even a low rate of deletion μ, changing (1,1) cells to (0,1) cells can greatly increase the relative proportion of (0,1) to (0,0) cells. To maintain the proportion of (0,1) relative to (0,0) cells,

$$p/(1-p) = \alpha\beta(e^{(t+(1-n-i)\nu)T} - 1)/(\mu(\alpha + \beta)). \quad (2)$$

Combining (1) and (2), and rearranging, leads to

$$e^{\alpha T}(\alpha e^{\beta T} + \beta e^{\alpha T})/((e^{\beta T} - e^{\alpha T})(e^{(t+(1-n-i)\nu)T} - 1)) = 2\alpha^3\beta^2/(i\nu\mu(\alpha + \beta)). \quad (3)$$

The 2 comes because only one half of the cells lacking the element at the site of interest at the start of environment 1 have the opportunity to have transpositions into the site of interest, and thus the effective transposition rate is half that in (1).

Because approximately one half of the cells lacking an element at the site of interest have no element at the neutral site, whereas almost all of the cells with an element at the site of interest have an element at the neutral site, the effective values for selection are $\alpha = s - 1.5\,(t + (1 - n - i)\nu)$, $\beta = s + 1.5(t + (1 - n - i)\nu)$.

Results and Discussion

Here we consider a model for the spread and maintenance of mobile DNA sequences in a clonal host population. We imagine that there are three types of location into which mobile genetic elements can insert: harmful sites (insertions into which are rapidly eliminated by natural selection), neutral sites, and what we call a site of interest, which we imagine is a structural gene encoding a product that is harmful in one environment and useful in another. (The harm does not have to be toxicity, but it could merely be the inefficient production of an unnecessary protein product.) Transposition occurs at a constant rate υ per element copy and, for each transposition event, the probabilities of insertion into the three types of sites are $(1 - n - i)$, n, and i, respectively. All transpositions are replicative. There is also a deletion rate of μ, where $\mu \ll \upsilon$. For simplicity, the model assumes all deletions to be precise excision events. The population cycles between two environments, for T generations in each. In environment 1, insertions into the site of interest are selectively favored, whereas, in environment 2, cells with no insertions into the site of interest are favored. Each of these selections operates at a rate s against the nonfavored genotype. Each element also imposes a fitness cost t, which is irrespective of its position.

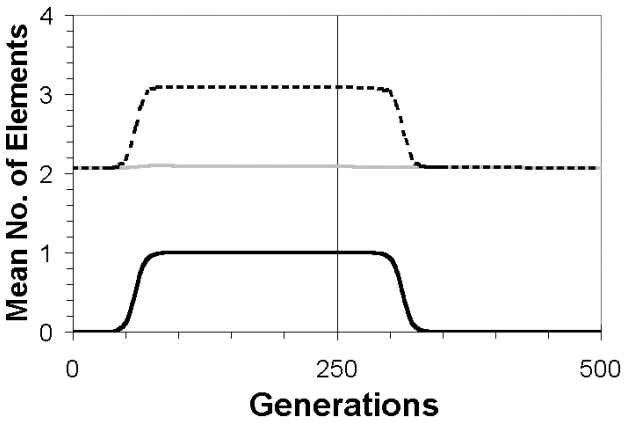

FIG. 1.—Changes in the mean number of transposable elements during an equilibrium cycle. $T = 250$, $n = 0.099$, $i = 0.001$, $t = 0.001$, $s = 0.2$, $\upsilon = 10^{-4}$ and $\mu = 10^{-7}$. Long-term maintenance of elements is through insertions at neutral sites (unbroken gray line). Mean element number at neutral sites receives a small "hitch-hiking" boost, as the insertion at the site of interest (unbroken black line) sweeps through the population at the start of environment 1 and then declines slowly throughout the rest of the cycle asa result of selection against elements of $(t + 0.9\upsilon)$. The broken black line is total mean element number.

The population is defined as a collection of cells that differ in the number of elements that they have at neutral sites and at the site of interest. The proportion of a given cell type changes each generation as a result of deletion, selection, and transposition. In the model there is the spread, during environment 1, of cells with an insertion at the site of interest. Then, during environment 2, there is selection in favor of cells that lack such insertions. Predictably, the frequency of cells with insertions at the site of interest cycles to track the changing environments. An example of an equilibrium cycle is shown in figure 1. We have explored the outcomes of such a model, concentrating in particular on the lengths of the cycles that allow the maintenance of the mobile elements in the population in an equilibrium cycle. Figure 2 shows the range of cycle lengths that permit the maintenance of the element for various parameter values.

The model has the outcome that the numbers of elements at the site of interest and the neutral sites are intimately connected. At the start of environment 1, typically no cells possess an element at the site of interest, and so creation of this mutation requires transpositional donor elements at neutral sites. Selection will favor the genotypes with a large number of elements at the neutral sites, because these genotypes will be quick to find an insertion into the site of interest. Linkage disequilibrium will then increase the frequency of neutral sites as selection acts at the site of interest. At the end of environment 1 there is fixation or near-fixation of cells with insertions in the site of interest, most of which also have insertions in neutral sites. Thus, despite the constant selection against them, elements at neutral sites are favored. The continued maintenance of elements at neutral sites arises not because of any foresight or anticipation of the need for such elements to act as donors of future insertions into selected sites. Rather, maintenance is a consequence of the linkage disequilibrium between

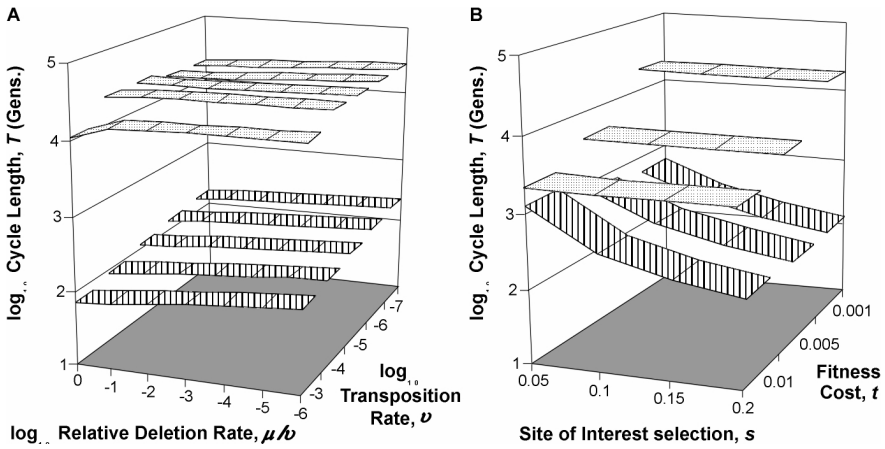

FIG. 2.—The maximum and minimum values for T that allow persistence of the elements. $n = 0.099$, $i = 0.001$. (A) Transposition, υ, and Deletion, μ, frequencies have minimal impact on T for an infinite population size. $t = 0.001$, $s = 0.2$. (B) Selection at the site of interest, s, has a large effect on minimum T. Selection against elements, t, has a large effect on maximum T. $\upsilon = 10^{-4}$ and $\mu = 10^{-7}$.

neutral and positive selected sites at the start of environment 1.

The maintenance of the elements under the model is quite robust to changes in the parameters. This is particularly true of the length of cycles, T, which permit maintenance. Parameter values have a greater effect on the mean number of elements at equilibrium. Figure 3 shows the equilibrium mean number of elements at the start of environment 1 under three different sets of parameters, with differing s and t values, across the range of T in which elements are maintained. None of the parameter sets in the models shown allow maintenance of the mobile element in the absence of favored insertions during environment 1. Further, the conditions for selfish spread are not met. Also, over the whole cycle, the geometrical mean fitness of cells with elements at the site of interest is less than that of cells without such elements. Thus the selection at the site of interest could not maintain the elements in the absence of elements at neutral sites.

Because the model includes very many classes of genotype, the changes in frequency of which are complex functions of the parameters and frequencies of other classes, analytical solutions of conditions for element maintenance are impossible. It is possible, however, to produce approximate analytical predictions of the range of cycle lengths that allow the elements to persist. Very rapid cycling, in which the frequency of the elements at the site of interest oscillates as rapid shifts occur between environments 1 and 2, cannot maintain the elements. The minimum cycle length is a value of T sufficiently great that the frequency of elements at the site of interest drops in environment 2 to a value of the same order of magnitude as the rate of transposition into the site of interest, or $i\upsilon$, and rises sufficiently during environment 1 for deletion to be a significant force in generating the cells without these insertions that will be favored in environment 2. Table 1 compares exact results from the simulations with predictions from the approximate analytical model for the minimum values of T that permit elements to persist.

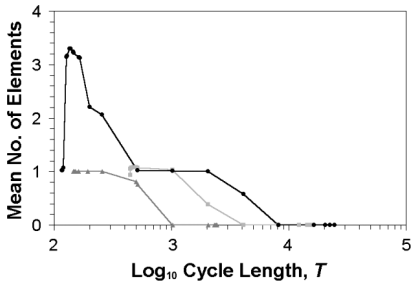

FIG. 3.—Mean numbers of elements per genome at the end of environment 2 in the equilibrium cycle as a function of $\log_{10}T$ for three different combinations of s and t. $s = 0.2$, $t = 0.001$ (black circles); $s = 0.2$, $t = 0.01$ (dark gray triangles); $s = 0.05$, $t = 0.01$ (light gray squares). s affects minimum T, t mainly affects maximum T, while both have an effect on mean element number.

The upper limit of T that allows maintenance of the elements is defined by the loss of all elements at rate of $t + (1 - n - i)\upsilon$ per element during environment 2. In the simulations, if T becomes too large, the proportion of cells creating transpositions at the start of environment 1 becomes too small to be recorded in the program. This is unrealistic. In any finite population, extinction of elements in environment 2 by drift will be expected for very much lower values of T than this, and the competing effects of base substitution mutations will become significant. If the time in environment 2 is too long, the elements will be lost, because selfish spread will be incapable of maintaining them. Only by horizontal transfer occurring during a future environment of type 1 could the elements be reintroduced.

The model can also be applied to situations in which a low level of horizontal transfer is allowed. Figure 4 shows the impact of a level of horizontal transfer of elements where none, half, or all of the transpositions enter a cell different from the donor cell. Again, selection against the element is still too strong for selfish spread to maintain elements. The introduction of horizontal transfer reduces mean element number. The effect is small at low levels of horizontal transfer but becomes severe as the level of horizontal transfer approaches 50%. This is because insertions into the site of interest that are made as

Table 1
Approximate and Exact Values of the Minimum Environment Length, *T*, Required for the Maintenance of Elements

	t					
s	0.01		0.005		0.001	
$\upsilon = 10^{-3}$ $\mu = 10^{-6}$						
0.05	**810**[a]	*1107*	**526**	610	**393**	*444*
0.10	**284**	*301*	**240**	249	**199**	*212*
0.15	**178**	*175*	**158**	156	**134**	*137*
0.20	**131**	*122*	**118**	112	**102**	*98*
$\upsilon = 10^{-4}$ $\mu = 10^{-7}$						
0.05	**931**	*1192*	**622**	*693*	**463**	*509*
0.10	**338**	*352*	**287**	*292*	**235**	*246*
0.15	**212**	*206*	**189**	*184*	**159**	*159*
0.20	**156**	*145*	**142**	*133*	**120**	*116*

[a] The number in bold is the accurate value and that in italics is that predicted from the approximate model as described in *Materials and Methods*. In each case, $n = 0.099$, $i = 0.001$.

a result of a horizontal transfer, rather than by transposition from a linked neutral site, do not give the frequency of neutral insertions a boost due to "hitchhiking." [An alternative model of horizontal gene transfer was also tested. This time, a proportion of the population replaced the site of interest with that from a random member of the population. The effects of gene transfer were similar to horizontal transmission of TEs, with an even greater reduction in mean element number (data not shown)].

The above frequency-based model applies to a population with an infinite population size. We have thus considered an equivalent model in which there are a finite number of individuals. Figure 5 shows simulations of the spread of the mobile sequence from a starting frequency of 0.5 at neutral sites for population sizes of 10^5 to 10^9, and transposition rates, υ, of 10^{-4} and 10^{-5} ($\mu = 0.001\upsilon$). The most significant aspect of these simulations is that a very high stochasticity is seen in what are, in terms of genetic drift, very large populations. The cause of this is the genetic bottlenecks encountered by the population at the start of each environmental phase, when only the very small subset of cells, those that either received (at the start of environment 1) or lost (at the start of environment 2) the element at or from the site of interest, will leave any descendants. Thus, the effective population size has to be orders of magnitude greater than the reciprocal of the transposition rate before the behavior of the population starts to approximate that of an infinite population.

It has been observed that the level of genetic variation in bacterial populations is less than that expected from their apparent population sizes (Maynard Smith 1991; Levin and Bergstrom 2000). This lack of variation is consistent with a reduction in effective population size that results from recurrent selective sweeps and periodic selection (Koch 1974; Levin 1981; Guttman 1997). Periodic selection is normally considered in terms of consecutive adaptations to a single environment (Atwood, Schneider, and Ryan 1951; Koch 1974). This model presents an alternative form of periodic selection, in which amino acid sequences will not have changed adaptively but neutral variation will remain lower than expected.

The model presented here involves some form of environmental cycling. There is no reason to expect the environmental cycles to be synchronized in remote locations. Rather we believe that the population structure of many bacteria will be a "metapopulation," with environmental fluctuations operating locally, maintaining mobile elements by this mechanism, but with low levels of gene flow between local populations. The gene flow has to be sufficiently low that the advantageous genotypes required at the starts of environment 1 and environment 2 are generated by transposition and deletion, respectively, rather than being introduced by migration from other populations.

Is this a likely model for bacterial response to fluctuating environments? First, do environments change in ways that create fitness advantages for insertion mutations? The introduction gives examples of advantageous insertions, inactivating genes in some environments. In addition, the phenomenon of mutations showing a growth advantage in stationary phase ("GASP") in *Escherichia coli* (Vulic and Kolter 2001), some of which are null mutations (Zinser and Kolter 2000), suggests one possible type of transient insertion mutation whose advantage is environment-dependent. Furthermore, we have shown (Edwards, Sockett, and Brookfield 2002) that, in experimental shaking cultures of *E. coli*, insertions creating null mutations in flagella genes show considerable

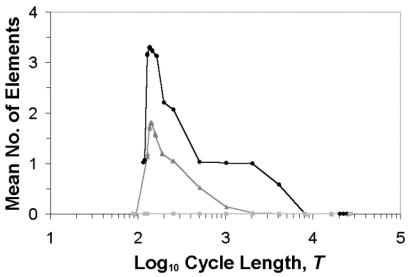

FIG. 4.—The impact of high levels of horizontal transfer on element maintenance: $h = 0$ (black circles); $h = 0.5$ (dark gray triangles); $h = 1$ (light gray squares). Although the range of parameter space is affected very little, the mean element number is substantially reduced as a result of the uncoupling of insertions at the site of interest from neutral insertions. $n = 0.099$, $i = 0.001$, $t = 0.001$, $s = 0.2$, $\upsilon = 10^{-4}$ and $\mu = 10^{-7}$.

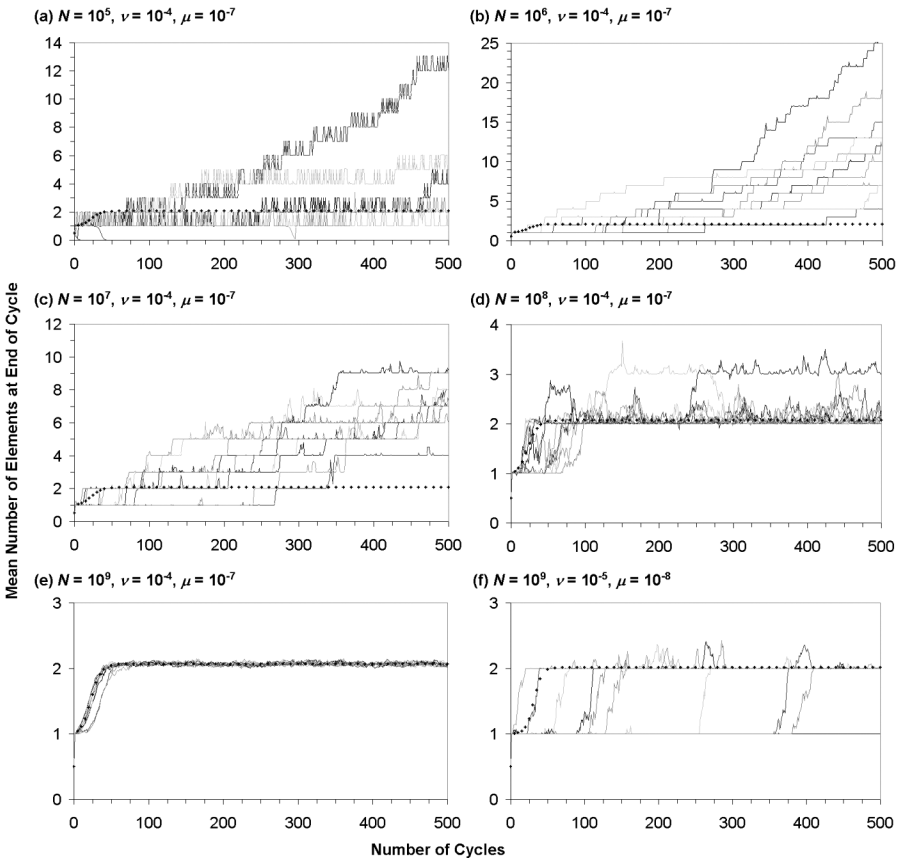

FIG. 5.—Simulations of the rise of elements from an initial frequency of 0.5 at neutral sites. In each case, ten replicate finite populations (solid lines) are compared to the infinite population expectation (+). $T = 250$, $s = 0.2$, $t = 0.001$, $n = 0.099$, $i = 0.001$. (a)–(e) As the population size increases from 10^5 to 10^9 ($\upsilon = 10^{-4}$, $\mu = 10^{-7}$), finite population simulations resemble the infinite expectation more closely. Note that the reduced variance in TE copy number reduces selection against elements, allowing them actually to fare better that the infinite expectation in smaller populations. (e)–(f) Keeping the total population size constant at 10^9, but decreasing transposition and deletion rates tenfold increases stochasticity by reducing the size of the bottleneck at each cycle.

(10%–20%) selective advantages over progenitor cells in which the pathway of flagella production is intact. One might expect that a regular cycle of environmental change would trigger the evolution of environmental sensing systems that would sequentially inactivate and reactivate the gene of interest in response to the cycles. Its greatly enhanced speed and efficiency would make the evolution of such a system of gene regulation probable if the population had a long history of exposure to two specific alternating environments. However, in our model, the gene of interest is restored to wild-type in the whole local population at the end of environment 2, and thus the model does not require that the gene of interest that is inactivated be the same in each cycle. Element maintenance in the model is also fairly robust to increases in the length of environment 2, and the regular periodicity of the cycling is not necessary. Thus, we have a model, not of a population in an environment that fluctuates regularly between two alternative states, but one in which a myriad of environments will successively select for insertions into, and reactivation of, different genes. For each transient beneficial insertion, elements will receive a boost in numbers, provided that the selection persists long enough for the insertion to reach fixation in the local population.

Mobile genetic elements are not the only source of mutation. In reality, base substitutions will also occur, and they may inactivate the gene of interest during environment 1. The frequency of such inactivating base substitutions may well approach the rate of insertion of elements into the gene of interest. If so, at the end of environment 1, when all cells have inactivating mutations in the site of interest, only about one half of the cells will have insertion mutations. However, during environment 2, the rate of excision of the element will be orders of magnitude higher than is the rate of back mutation to revert a base substitution mutation. Thus, at the end of environment 2, the proportion of cells that will have gone through the transposition-excision cycle will be orders of magnitude higher than the proportion of cells that have gone through a base substitution-reversion cycle, and the power of the former to maintain the elements will be little affected by the possibility of the latter.

Of course, mobile elements may be lost through processes other than precise excision, including imprecise excision or lack of function mutations. This will serve to increase the effective rate of deletion from neutral sites, and to reduce the maximum length T in which elements are maintained. The dynamics of elements at the site of interest

would not be greatly affected, as only cells experiencing a precise excision here have any long-term future.

The model potentially applies to all bacterial species. In addition, it may have some relevance for other clonal groups, including some plants and animals, such as the Bdelloid rotifers, which are an anciently clonal group of animals. These have been found to lack retrotransposons (Archipova and Meselson 2000). The absence of retrotransposons would be consistent with the absence of a precise deletion process for these elements, and their inability to be maintained in a clonal host by the mechanism that we model here. The model explains the maintenance of mobile elements not by sex, but by selection for their transiently advantageous effects. Whether bacteria are effectively sexual depends on bacterial species studied (Maynard Smith et al. 1993; Guttman 1997; Maynard Smith and Smith 1998). In some species, there is little sign of the linkage disequilibrium expected under clonality. It is important to remember that the level of sex and recombination required to create linkage equilibrium is not the same as that required to allow transposable elements to persist. *Escherichia coli*, for example, despite manifesting strong evidence of some horizontal gene-transfer events with important evolutionary consequences, has a very low level of recombination per generation (Milkman and Bridges 1990; Guttman and Dykhuizen 1994). To be effectively sexual in terms of linkage disequilibria for neutral genes, the time to common ancestry of randomly chosen individuals must be ancient compared to the reciprocal of the rate of recombination between lineages. The time to common ancestry will be the reciprocal of the effective population size. For the sexual spread of transposable elements, however, the rate of sexuality has to be large relative to the strength of selection against the elements (which we model in the range from 0.001 to 0.01). A level of sex and recombination which is enough to create effective sexuality in terms of linkage equilibrium will be insufficient to allow selfish elements to spread by sex in the face of selection.

The common feature of transposable elements is their mobility, and this is the only attribute required by the model. Here, we have explicitly modeled unregulated replicative transposition. Regulation of transposition rates such that the number of transpositions per generation is independent of the number of elements in the cell (above zero) has the predictable affect of reducing TE number (not shown). As long as a cell has at least one element, there is no longer an advantage for having more copies. The range of parameter space that can maintain elements is largely unchanged because there is generally only one neutral insertion retained at the extremes of T, and so regulation of transposition here has no effect. Some bacterial transposable elements may increase rates of transposition (Kleckner 1981; Blot 1994; Mahillon and Chandler 1998) or precise excision (Aleshkin, Kadzhaev, and Markov 1998) in times of physiological stress. This has not been modeled but will intuitively benefit elements and the lower limit of their range by accelerating the time taken for advantageous insertions to spread through the population at the start of environment 1.

We have presented here a general model for the maintenance of mobile genetic elements through their ubiquitous capacity to generate insertion mutations. If such mutations experience transient positive selection, then linkage disequilibrium with neutral donor elements is sufficient to increase mean element number. A fluctuating environment of this kind would generate periodic selection but would not necessarily leave a legacy of amino acid substitutions as is seen following successive selective sweeps adapting to a stable environment. In direct contrast to the selfish DNA hypothesis, horizontal transfer weakens the maintenance of elements in this system.

Acknowledgments

This work was funded by the University of Nottingham. We thank two anonymous referees for helpful comments on the first draft of the manuscript.

Literature Cited

Aleshkin, G. I., K. V. Kadzhaev, and A. P. Markov. 1998. High and low UV-dose responses in SOS-induction of the precise excision of transposons Tn*1*, Tn*5* and Tn*10* in *Escherichia coli*. Mutat. Res. **401**:179–191.

Allgood, N. D., and T. J. Silhavy. 1988. Illegitimate recombination in bacteria. Pp. 309–330 in R. Kucherlapati and G. R. Smith, eds. Genetic recombination. American Society of Microbiology, Washington, D.C.

Archipova, I., and M. Meselson. 2000. Transposable elements in sexual and ancient asexual taxa. Proc. Natl. Acad. Sci. USA **97**:14473–14477.

Atwood, K. C., L. K. Schneider, and F. J. Ryan. 1951. Periodic selection in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **37**:146–155.

Blot, M. 1994. Transposable elements and adaptation of host bacteria. Genetica **93**:5–12.

Cavalier-Smith, T. 1980. How selfish is DNA? Nature **285**:617–618.

Chao, L., and E. C. Cox. 1983. Competition between high and low mutating strains of *Escherichia coli*. Evolution **37**:125–134.

Chao, L., and S. M. McBroom. 1985. Evolution of transposable elements: an IS*10* insertion increases fitness in *Escherichia coli*. Mol. Biol. Evol. **2**:359–369.

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

Condit, R., F. M. Stewart, and B. R. Levin. 1988. The population biology of bacterial transposons: *a priori* conditions for maintenance as parasitic DNA. Am. Nat. **132**:129–147.

Edwards, R. J., R. E. Sockett, and J. F. Y. Brookfield. 2002. A simple method for genome-wide screening for advantageous insertions of mobile DNAs in *Escherichia coli*. Current Biol. **12**:863–867.

Guttman, D. S. 1997. Recombination and clonality in natural populations of *Escherichia coli*. Trends Ecol. Evol. **12**:16–22.

Guttman, D. S., and D. E. Dykhuizen. 1994. Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. Science **266**:1380–1383.

Hammerschidt, S., R. Hilse, J. P. M. van Putten, R. Gerardy-Schahn, A. Unkenmeir, and M. Frosch. 1996. Modulation of cell surface sialic acid expression in *Neisseria meningitidis* via a transposable genetic element. EMBO J. **15**:192–196.

Hickey, D. A. 1982. Selfish DNA: a sexually-transmitted nuclear parasite. Genetics **101**:519–531.

Kleckner, N. 1981. Transposable elements in prokaryotes. Annu. Rev. Genet. **15**:341–404.

Koch, A. L. 1974. The pertinence of the periodic selection phenomenon to prokaryote evolution. Genetics **77**:127–142.

Levin, B. R. 1981. Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. Genetics **99**:1–23.

Levin, B. R., and C. T. Bergstrom. 2000. Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes. Proc. Natl. Acad. Sci. USA **97**:6981–6985.

Mahillon, J., and M. Chandler. 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. **62**:725–774.

Mao, E. F., L. Lane, J. Lee, and J. H. Miller. 1997. Proliferation of mutators in a cell population. J. Bacteriol. **179**:417–422.

Maynard Smith, J. 1991. The population genetics of bacteria. Proc. R. Soc. Lond. Ser. B **245**:37–41.

Maynard Smith, J., and N. H. Smith. 1998. Detecting recombination from gene trees. Mol. Biol. Evol. **15**:590–599.

Maynard Smith, J., N. H. Smith, M. O'Rourke, and B. G. Spratt. 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. USA **90**:4384–4388.

Milkman, R., and M. M. Bridges. 1990. Molecular evolution of the *Escherichia coli* chromosome. III. Clonal frames. Genetics **126**:505–517.

Mills, J. A., M. M. Venkatesan, L. S. Baron, and J. M. Buysse. 1992. Spontaneous insertion of an IS*1*-like element into the *virF* gene is responsible for avirulence in opaque colonial variants of *Shigella flexneri* 2a. Infect. Immun. **60**:175–182.

Modi, R. I., L. H. Castilla, S. Puskas-Rozsa, R. B. Helling, and J. Adams. 1992. Genetic changes accompanying increased fitness in evolving populations of *Escherichia coli*. Genetics **130**:241–249.

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.

Taddei, F., M. Radman, J. Maynard-Smith, B. Toupance, P. H. Gouyon, and B. Godelle. 1997. Role of mutator alleles in adaptive evolution. Nature (Lond.) **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.

Vulic, M., and R. Kolter. 2001. Evolutionary cheating in *Escherichia coli* stationary phase cultures. Genetics **158**:519–526.

Ziebuhr, W., V. Krimmer, S. Rachid, I. Lossner, F. Gotz, and J. Hacher. 1999. A novel mechanism of phase variation of virulence in *Staphylococcus epidermidis*: evidence for control of the polysaccharide intercellular adhesin synthesis by alternating insertion and excision of the insertion sequence element IS*256*. Mol. Microbiol. **32**:345–356.

Zinser, E. R., and R. Kolter. 2000. Prolonged stationary-phase incubation selects for lrp mutations in *Escherichia coli* K-12. J. Bacteriol. **182**:4361–4365.

Jeffrey Long, Associate Editor

Accepted August 29, 2002