INTRODUCTION

Natural genetic transformation can facilitate gene transfer in many genera of bacteria and requires the presence of extracellular DNA. Although cell lysis can contribute to this extracellular DNA pool, several studies have suggested that the secretion of DNA from living bacteria may also provide genetic material for transformation. This paper reviews the evidence for specific secretion of DNA from intact bacteria into the extracellular environment and examines this behaviour from a population-genetics perspective. A mathematical model demonstrates that the joint action of DNA secretion and transformation creates a novel type of gene-level natural selection. This model demonstrates that gene-level selection could explain the existence of DNA secretion mechanisms that provide no benefit to individual cells or populations of bacteria. Additionally, the model predicts that any trait affecting DNA secretion will experience selection at the gene level in a transforming population. This analysis confirms that the secretion of DNA from intact bacterial cells is fully explicable with evolutionary theory, and reveals a novel mechanism for bacterial evolution.

DNA secretion and gene-level selection in bacteria

Jeremy A. Draghi and Paul E. Turner

Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA

Natural genetic transformation can facilitate gene transfer in many genera of bacteria and requires the presence of extracellular DNA. Although cell lysis can contribute to this extracellular DNA pool, several studies have suggested that the secretion of DNA from living bacteria may also provide genetic material for transformation. This paper reviews the evidence for specific secretion of DNA from intact bacteria into the extracellular environment and examines this behaviour from a population-genetics perspective. A mathematical model demonstrates that the joint action of DNA secretion and transformation creates a novel type of gene-level natural selection. This model demonstrates that gene-level selection could explain the existence of DNA secretion mechanisms that provide no benefit to individual cells or populations of bacteria. Additionally, the model predicts that any trait affecting DNA secretion will experience selection at the gene level in a transforming population. This analysis confirms that the secretion of DNA from intact bacterial cells is fully explicable with evolutionary theory, and reveals a novel mechanism for bacterial evolution.
through lysis. Transformation via contact between intact donor cells and recipients was found to produce over a thousand times as many transformants as did transformation with an equal quantity of purified DNA (Stewart et al., 1983). In the same experiments, transformation with intact donor cells also produced over an order of magnitude more transformants than did transformation with saturating concentrations of purified DNA. Cell death did not occur at detectable levels in these conditions, and reciprocal experiments established that each of two strains was equally able to donate and receive DNA. These results suggest that in this strain and environment, the release of DNA by intact cells, not lysis or a cryptic infectious agent, provides the primary material for transformation.

A detailed study of the temporal progression of competence in B. subtilis 168 illustrates the distinction between secreted DNA and DNA released by cell lysis (Lorenz et al., 1991). When this strain was cultured under conditions optimal for competence, extracellular DNA concentrations peaked briefly in late exponential phase, declined to nearly zero, then slowly increased through stationary phase. This initial peak is exactly coincident with an equally brief period of competence in the culture: during this same period cell death is near the minimum measured during the growth of the culture. In contrast, the later gradual increase of extracellular DNA tracks levels of cell lysis, as measured by extracellular catalase activity.

In addition to these lines of indirect evidence for secretion in several naturally transforming species, a genetic island found in 80% of N. meningitidis genomes was recently determined to confer active DNA secretion through a type IV secretion system (Hamilton et al., 2005). Directed mutation resulted in several mutants that produced much less extracellular DNA without a measurable decrease in the rate of lysis. This discovery suggests that genomic approaches may be the key to uncovering mechanisms of secretion in other prokaryotes, and may provide a basis for an evolutionary understanding of the phenotype.

The above studies support the idea that DNA secretion evolved at least once in bacteria and is selectively favoured in at least some environments. But how could the secretion of expensive DNA molecules possibly be adaptive? One possibility is that DNA secretion is selected by the benefits it provides to populations of bacteria: as a complement to transformation, secretion could facilitate an efficient exchange of alleles among conspecifics (Stewart & Carlson, 1986). However, such a strategy appears patently unstable: mutants which did not secrete DNA would reap any benefits while avoiding the costs, and would quickly drive secreting cells toward extinction. Although group selection may influence DNA secretion in some populations, here we propose an explanation for secretion at a much lower level of organization. The mathematical model presented below argues that secretion of genetic material in a transforming population can lead to a novel variant of gene-level natural selection. This form of gene-level selection, which we call selective conversion, acts independently of selection on organisms and populations and could maintain the DNA secretion phenotype in bacterial populations.

**THEORY AND RESULTS**

To evaluate whether gene-level selection is a plausible explanation for DNA secretion, we derived a quantitative description of the dynamics of alleles at a locus L controlling DNA secretion. This locus can be polygenic, and has two alleles, S and W. W, the wild-type, has relative fitness 1 and does not secrete DNA. S, the DNA-secreting mutant, has relative fitness \(1 - c\), where \(c\) is the cost of secretion to growth, and secretes copies of locus L at the rate \(s\). Define \(p\) as the frequency of the S allele in the population. We assume the population size is stable, as in a turbidostat, so cells die at rate \(r\). Cells divide to replace losses in proportion to their fitness. Based on these definitions, we can derive the change in \(p\), the frequency of genotype S, as a differential equation:

\[
\frac{dp}{dt} = \Delta G + \Delta T
\]

where \(\Delta G\) represents the change due to differential growth, and \(\Delta T\) the change caused by transformation and secretion. First, we can derive \(\Delta G\): cells of genotype S are lost at a rate \(pr\), and are created at a rate equal to their fitness, divided by the mean fitness of the population, multiplied by \(r\):

\[
\frac{dp}{dt} = \frac{r - c p s (1 - c)}{p (1 - c) + (1 - p)} - r p + \Delta T
\]

\[
= \frac{-r c p (1 - p)}{1 - p c} + \Delta T
\]

Of course, if the cost is non-zero, equation 2 shows that the frequency of genotype S declines to zero when only growth is considered. To complete this description of allele dynamics, we must add \(\Delta T\), the effects of transformation. Define \(t\) as the rate at which extracellular DNA surrounding a competent cell is taken up; the rate at which a cell is transformed is then the product of \(r\) and the number of copies of locus L in the extracellular environment. Define \(D\) as this number, and \(d\) as the rate at which this DNA is degraded under extracellular conditions. Finally, assume that cell death also contributes to the pool of extracellular DNA, and that cells are tightly packed, such that any piece of extracellular DNA is close enough to at least one cell to be taken up by transformation. This last assumption ensures that, in a well-mixed population, each cell has access to the DNA released from, on average, one other cell. We can then describe the effect of these four factors, cell death, secretion, degradation, and transformation, on \(D\) as follows:

\[
\frac{dD}{dt} = r + s p - d \cdot D - t \cdot D
\]

To determine the equilibrium value of \(D\), we simply set \(dD/dt\) to zero and solve equation 3, yielding:
Equation 4, multiplied by \( t \), is the transformation rate at locus L; to determine the effects of this transformation on \( p \), we must calculate the fraction of transformation events which result in each genotype. As above, we simply sum the rates of loss and gain: genotype S cells are lost to transformation at a rate \( p \cdot r \cdot D \), and gained at a rate \( t \cdot D \), multiplied by the fraction of extracellular DNA that carries the S allele:

\[
\Delta T = t \cdot D \left( \frac{p \cdot (s + r)}{p \cdot s + r} - p \right)
\]  

Combining this with equations 2 and 4 yields:

\[
\frac{dp}{dt} = t \cdot p \cdot s + r \cdot \left( \frac{p \cdot (s + r)}{p \cdot s + r} - p \right) - \frac{r \cdot c \cdot p \cdot (1 - p)}{1 - c \cdot p}
\]  

To determine if one allele will drive the other to extinction, or if both will coexist, we can find equilibrium solutions by setting equation 6 equal to zero. We can then solve the resulting equation for any variable, such as for \( c \):

\[
c = \frac{s \cdot t}{d \cdot r + r \cdot t + p \cdot s \cdot t}
\]

Fig. 1 plots values of \( c \) calculated from equation 7; each plot shows how the system behaves for particular costs when three other parameters are held constant and the remaining parameter is varied. In each plot, two curves, the upper plotted for \( p = 0.01 \), the lower for \( p = 0.99 \), define three regions of parameter space. The white region contains pairs of costs \( (c) \) and secretion rates \( (s) \) at which genotype S cells make up less than 1% of the population at equilibrium. Similarly, the dark grey region contains pairs of \( c \) and \( s \) values at which genotype S cells make up more than 99% of the population. For pairs of values in the light grey region separating these areas there is an equilibrium value of \( p \) between 0.01 and 0.99, suggesting the possibility of a stable polymorphism at these values. Analysis of the derivative of equation 6 at these equilibrium points indicates that they are stable for the range of conditions in Fig. 1 (results not shown), confirming coexistence of the two genotypes for these parameter values.

**Fig. 1.** Equilibrium states for equation 6. In each plot, the top curve represents points where the allele frequencies are stable for \( p = 0.01 \), the bottom curve corresponds to stability at \( p = 0.99 \), and the light grey area bounded by the curves is the range of values for which the population is stably polymorphic. The white region represents parameter values for which the W genotype is fixed, and the dark grey region contains values at which genotype S cells make up more than 99% of the population. For pairs of values in the light grey region separating these areas there is an equilibrium value of \( p \) between 0.01 and 0.99, suggesting the possibility of a stable polymorphism at these values. Analysis of the derivative of equation 6 at these equilibrium points indicates that they are stable for the range of conditions in Fig. 1 (results not shown), confirming coexistence of the two genotypes for these parameter values.

\( D = \frac{r + s \cdot p}{d + t} \)  

\[
D_{\text{rzd}}(4) \times t \text{, the transformation rate at locus L; to determine the effects of this transformation on } p, \text{ we must calculate the fraction of transformation events which result in each genotype. As above, we simply sum the rates of loss and gain: genotype S cells are lost to transformation at a rate } p \cdot r \cdot D, \text{ and gained at a rate } t \cdot D, \text{ multiplied by the fraction of extracellular DNA that carries the S allele:}
\[
\Delta T = t \cdot D \left( \frac{p \cdot (s + r)}{p \cdot s + r} - p \right)
\]

Combining this with equations 2 and 4 yields:

\[
\frac{dp}{dt} = t \cdot p \cdot s + r \cdot \left( \frac{p \cdot (s + r)}{p \cdot s + r} - p \right) - \frac{r \cdot c \cdot p \cdot (1 - p)}{1 - c \cdot p}
\]

To determine if one allele will drive the other to extinction, or if both will coexist, we can find equilibrium solutions by setting equation 6 equal to zero. We can then solve the resulting equation for any variable, such as for \( c \):

\[
c = \frac{s \cdot t}{d \cdot r + r \cdot t + p \cdot s \cdot t}
\]
DISCUSSION

The results in Fig. 1 demonstrate that in our model, genotypes that secrete DNA can invade, and sometimes replace, wild-type non-secretors under certain conditions. Increasing the rate of transformation or secretion, or decreasing the rate at which extracellular DNA degrades, allows secretion to succeed at higher costs. Conversely, increasing the growth rate of the population decreases the maximum cost at which secretion is still evolutionarily favoured. These results can be easily understood by interpreting equation 6 as describing selection at two levels, each with its own strength and tempo. One level acts on the growth of individual cells with strength $c$, at a speed determined by the rate of population turnover, $r$. A second level acts on alleles mobilized by secretion at a rate $s$, and integrated by transformation, at a pace jointly set by the rate of transformation, $t$, and the rate of DNA degradation, $d$. We call the second influence on allele frequencies selective conversion to emphasize both its mathematical congruence with natural selection and the interconversion of alleles in genomes by transformation.

Our results, indicating that the balance of these two levels of selection can favour DNA secretion, suggest new emphases for the empirical study of DNA secretion. Uncovering the genetic bases for DNA secretion will facilitate the study of knockout mutations: these mutants will allow $s$ and $c$ to be measured in defined environments, directly testing the predictions of our model. Another significant empirical question is the content of secreted DNA: secretion of the whole genome, a single locus, or any intermediate possibility are all valid phenotypes within the model. Although we have tacitly assumed secretion of the entire genome, locus-specific secretion would likely present lower costs to the cell while retaining the same effective $s$. In fact, the apparent benefits of a specific secretion system suggest a critical test of our hypothesis: if upon further study most secretion systems are found to be non-specific, then selection for secretion solely on the gene level would probably be an incomplete explanation.

The results above also describe the ideal conditions for DNA secretion to exist and for selective conversion to be relevant: closely packed cells with minimal turnover and high rates of DNA uptake, in an environment supporting high levels of metabolic activity and long persistence of extracellular DNA. An interesting ecological question is whether these conditions are consistent with the dense, spatially structured populations of bacterial biofilms. Gene exchange through transformation has been observed in several monoculture laboratory biofilms, often at transfer rates exceeding those in liquid cultures by orders of magnitude (Li et al., 2001; Hendrickx et al., 2003; Molin & Tolker-Nielsen, 2003). These rates, however, decline with biofilm age. Similarly, growth rates in monoculture biofilm development can significantly exceed those in similar planktonic conditions but also decrease with biofilm age (Li et al., 2001). Though these laboratory microcosms suggest that rapid rates of transformation and slow rates of growth may plausibly co-occur, the balance of these rates in natural biofilms composed of many species, growing in complex and variable environments, is completely unknown. The theoretical model presented here underlines the importance of measuring transformation rates in natural biofilms to better understand the processes of adaptation that shape the populations within them.

To fully evaluate the application of our model, we must also consider how violations of the underlying assumptions affect our conclusions. Several important assumptions relate to the pool of extracellular DNA. In our model, extracellular DNA is generated only by the death of and secretion by conspecific cells. Other sources may contribute DNA which is not homologous to locus $L$, but which may compete with relevant DNA for uptake: such a situation is easily modelled as a decrease in $t$, with a corresponding decrease in selection for secretion. We have also assumed that cells are tightly packed, such that extracellular DNA is always adjacent to a cell. Deviations from this assumption may be significant in sparse populations or populations of fluctuating size. Finally, we have assumed that every homologous piece of DNA which enters a cell transforms the genome: in species with restriction enzymes or other nuclease systems, this assumption may overestimate the true rate of transformation. Degradation of DNA within the cell could be modelled by multiplying the first term in equation 6 by the fraction of intracellular copies of locus $L$ which successfully transform the cell.

Other assumptions relate to the idealized growth conditions of our model. For mathematical simplicity, our model considers population turnover to be independent of the frequency of the $S$ genotype; if we instead allow the death rate to depend upon fitness, and therefore upon $p$, the results are nearly identical (data not shown). Also, we have modelled transformation as continuous: however, many bacteria become competent only under particular conditions and densities. If transformation occurred at discrete intervals, there would be quantitative deviations from our continuous model, much as continuous models of growth deviate from the actual discrete process of growth. If generation times and intervals between transformation events differed greatly for a particular population, an approach based on difference equations would be necessary for precise quantitative prediction.

One other assumption worth considering is that transformation itself does not evolve. The same gene-level selection argument we have applied to a hypothetical secretion locus applies equally to any loci involved with transformation: if an allele which decreases the rate of transformation enters a population, then cells carrying that allele will tend to be transformed less often. All else being equal, this asymmetry will result in a small selective pressure against transformation. Although the selective forces acting on transformation are outside the scope of this paper, this point does compound the presumed costs of maintaining a...
transformation system, and highlights the apparent existence of selection in favour of transformation in the natural world.

Although the process we have called selective conversion has not to our knowledge been previously described, it does bear a resemblance to another form of gene-level selection: biased gene conversion in eukaryotes (Lamb & Helmi, 1982; Marais, 2003). Several studies have illustrated how biased gene conversion acts as a form of gene-level selection between homologous alleles (Lamb & Helmi, 1982; Nagylaki, 1983): typically small biases in conversion direction and the coupling of conversion rates to reproductive rates make this selective effect generally weak (Nagylaki, 1983). Selective conversion also resembles the evolutionary forces acting on genetic elements, such as transposons and plasmids, which can be transmitted both vertically and horizontally. Formally, models of plasmid population dynamics have much in common with our model: in both, the genetic entities promote their own transfer to other cells while reducing the fitness of their hosts. However, plasmids can be lost by segregation, which has no real analogue in our model, and rates of plasmid transfer are considered to be too low to allow plasmids without any benefit to the cell to persist (Bergstrom et al., 2000): these differences necessitate our novel approach to the dynamics of DNA secretion phenotypes.

Although our model is intended to account for the existence of DNA secretion, its results extend beyond the few genes directly responsible for secreting DNA. Any locus which affects the rate of DNA secretion, or the persistence and availability for transformation of extracellular DNA, will be subject to selective conversion in a transforming population. Since experimental evidence indicates that secretion is tied to DNA synthesis (Hara & Ueda, 1981; Stewart et al., 1983) and to some aspects of competence (Bulygina & Prozorov, 1973), the number of such loci could be great. Genes which affect both fitness and secretion will therefore experience selection at two levels, and the standard framework of fitness evolution may be inadequate for understanding their evolution.

It has been suggested that an appreciation for life in microbial communities will reveal selection and organization at levels above the individual cell (Shapiro, 1998; Teske & Stahl, 2002). In this case, however, examination of the community life of bacteria has uncovered selection at a lower level than the individual. Transformation is often distinguished from the other processes of horizontal gene transfer, conjugation and transduction, because the latter processes are mediated by infectious agents (Lorenz & Wackernagel, 1994): this analysis illustrates how genetic transformation further blurs the line between chromosomal gene and infectious agent. Selective conversion also illustrates how dramatically the processes of adaptation in bacteria may deviate from the picture of evolution developed for eukaryotes, and highlights the challenges and opportunities bacterial communities present to ecol ogists and evolutionary biologists.

ACKNOWLEDGEMENTS

Thanks to members of the Turner lab for helpful discussions. Jeremy Dragi was supported by a National Aeronautics and Space Administration Graduate Research Fellowship, and Paul Turner was supported by grant no. DEB-04-52163 from the National Science Foundation.

REFERENCES


