Bacterial Species and Speciation

FREDERICK M. COHAN

Department of Biology, Wesleyan University, Middletown, Connecticut 06459-0170, USA; E-mail: fcohan@wesleyan.edu

Abstract.—Bacteria are profoundly different from eukaryotes in their patterns of genetic exchange. Nevertheless, ecological diversity is organized in the same way across all of life: individual organisms fall into more or less discrete clusters on the basis of their phenotypic, ecological, and DNA sequence characteristics. Each sequence cluster in the bacterial world appears to correspond to an "ecotype," defined as a population of cells in the same ecological niche, which would all be out-competed by any adaptive mutant coming from the population. Ecotypes, so defined, share many of the dynamic properties attributed to eukaryotic species: genetic diversity within an ecotype is limited by a force of cohesion (in this case, periodic selection); different ecotypes are free to diverge without constraint from one another; and ecotypes are ecologically distinct. Also, ecotypes can be discovered and classified as DNA sequence clusters, even when we are ignorant of their ecology. Owing to the rarity and promiscuity of bacterial genetic exchange, speciation in the bacterial world is expected to be much less constrained than in the world of animals and plants. [Clonal structure; ecotype; periodic selection; sexual isolation; species concept.]

Wherever there is life, there are bacteria (Margulis and Schwartz, 1998; Madigan et al., 1999). Free-living bacteria are found in every environment that supports eukaryotes (Madigan et al., 1999). Also, no metazoan or metaphyte is known to be free of bacterial commensals or pathogens (Madigan et al., 1999). The most physically extreme habitats capable of supporting life can support extremophilic bacteria (e.g., Brock, 1978; Madigan and Marrs, 1997). Finally, anoxic habitats can support a great diversity of bacteria whose physiology does not require oxygen (Fenchel and Finlay, 1995). There is clearly tremendous ecological diversity within the prokaryotic world.

The ecological diversity among prokaryotes is patterned in much the same way as in eukaryotes: Individual organisms fall into discrete clusters on the basis of their phenotypic, ecological, and DNA sequence characteristics (Sneath, 1977; Broom and Sneath, 1981; Bridge and Sneath, 1983; Sneath and Stevens, 1985; Barrett and Sneath, 1994; Embley and Stackebrandt, 1997; Goodfellow et al., 1997). For example, Barrett and Sneath (1994) found that 315 strains from the Neisseriaceae (the family including the pathogens that cause gonorrhea and meningitis) fall into 31 distinct clusters on the basis of metabolic and other phenotypic characteristics. Likewise, sequence-based surveys of bacterial diversity yield discrete clusters in sequence space (Palys et al., 1997), as is also seen in eukaryotes (Mallet, 1995).

In the world of the highly sexual eukaryotes (including most plants and animals), the clustered pattern of diversity has long provided the primary basis for demarcating species in practice. For many decades, species have been demarcated as phenotypic clusters (Sokal and Crovello, 1970), and more recently, sequence data have allowed species to be demarcated as clusters in genotypic or sequence space (Mallet, 1995).

Species demarcation in bacteria has relied exclusively on the phenotypic and genotypic clustering of bacterial strains (Goodfellow et al., 1997). Bacterial species have long been understood to be phenotypic clusters (based largely on metabolic characters; Sneath, 1977). Numerical taxonomy has clearly demonstrated that bacterial strains fall into discrete phenotypic clusters and has provided a means for discovering species and for classifying unknown strains into already characterized species (Broom and Sneath, 1981; Bridge and Sneath, 1983).

Genotypic clustering has largely replaced phenotypic clustering as a primary criterion for demarcating bacterial species (Wayne et al., 1987). For decades, clustering based on whole-genome DNA–DNA hybridization between pairs of strains has served to distinguish species (Johnson, 1973; Wayne et al., 1987). More recently, systematists are using clustering of 16S rRNA (Stackebrandt and Goebel, 1994) and protein-coding (Palys et al., 1997, 2000) gene sequences as criteria for species demarcation.

Although clustering provides a taxonomic tool for classifying organisms into species, systematists and evolutionary biologists have widely believed that species are more than just clusters of very closely related and very similar organisms. Species are believed to have some fundamental dynamic properties as well. For instance, different species are thought to exploit different ecological niches (e.g., Ecological Species Concept; Van Valen, 1976). Also, genetic divergence within a species is thought to be constrained by one or more forces of cohesion (Cohesion Species Concept; Meglitsch, 1954; Templeton, 1989), most frequently genetic exchange (Biological Species Concept; Mayr, 1982). Different species are thought to have separate evolutionary fates, in that they are free to diverge without constraint from one another (Evolutionary Species Concept; Simpson, 1961; Wiley, 1978). In the world of plants and animals, one species is thought to split into two when the two groups have diverged sufficiently in their reproductive and ecological characteristics (e.g., Eldredge, 1985): The reproductive divergence prevents genetic exchange from reversing the genetic divergence between incipient species (Mayr, 1982), and the ecological divergence is necessary to allow competitive coexistence between the incipient species (Gause, 1934; May, 1973).

Here I address whether clusters observed in the bacterial world share the dynamic properties attributed to eukaryote species, and whether the mechanisms driving the origins of new species in bacteria might be shared with the eukaryotes. I will demonstrate that, despite basic differences in the nature of genetic exchange between bacteria and eukaryotes, bacterial species share many of the fundamental properties of eukaryotic species. Moreover, bacterial species can be demarcated in practice by the same sequence-cluster criteria used in eukaryotic systematics.

GENETIC EXCHANGE IN PROKARYOTES

The character of genetic exchange in prokaryotes differs profoundly from that in the most highly sexual eukaryotes, in several ways. First, recombination in bacteria is extremely rare. Several laboratories have taken a retrospective approach to determining the historical rates of recombination in nature. Based on surveys of diversity in allozymes, restriction-recognition sites, and gene sequences, recombination rates have been estimated from the degree of association between genes or between parts of genes (Hudson and Kaplan, 1985; Hudson, 1987; Hey and Wakeley, 1997). This approach has shown that a given gene segment is usually involved in recombination at about the same rate at which it is mutated, or less (Selander and Musser, 1990; Maynard Smith et al., 1993; Whittam and Ake, 1993; Roberts and Cohan, 1995).

Second, recombination in bacteria is much more promiscuous than is the case for plants and animals. Bacteria do not exchange genes as often as animals and plants, but when they do they are not nearly as fussy about their choice of partners. Animal groups typically lose the ability to exchange genes entirely by the time their mitochondrial DNA sequences are 3% divergent (Avise, 2000), although some animal subspecies as divergent as 16% can exchange genes in nature (Moritz et al., 1992). Bacteria, in contrast, can undergo homologous recombination with organisms as divergent as 25% in DNA sequence (and possibly more) (Duncan et al., 1989; Roberts and Cohan, 1993; Zawadzki et al., 1995; Vulic et al., 1997). There are, nevertheless, some important constraints on bacterial genetic exchange. Recombination that depends on vectors, such as bacteriophage-mediated transduction or plasmid-mediated conjugation, is limited by the host ranges of the respective vectors. Also, restriction endonuclease activity can greatly reduce the rate of recombination (Edwards et al., 1999; Milkman et al., 1999), although not in the case of recombination by transformation (Trautner et al., 1974; Cohan et al., 1991). Finally, homologous recombination is limited by the resistance to integration of divergent DNA sequences, because mismatch repair tends to reverse integration of a mismatched donorrecipient heteroduplex (Rayssiguier et al., 1989; Vulic et al., 1997) and because integration requires a 20–30-bp stretch of nearly perfectly matched DNA (Hsieh et al., 1992; Rao et al., 1995; Majewski and Cohan, 1998, 1999a). In both Gram-positive and Gramnegative bacteria, the rate of recombination decays exponentially with donor-recipient sequence divergence (Roberts and Cohan,

1993; Zawadzki et al., 1995; Vulic et al., 1997; Majewski and Cohan, 1998, 1999a; Majewski et al., 2000).

Third, recombination events in bacteria are highly localized to a small fraction of the genome. The sizes of segments transduced or transformed in the laboratory are frequently less than several kilobases (McKane and Milkman, 1995; Zawadzki and Cohan, 1995). Surveys of sequences in nature support the conclusion that each recombination event is highly localized within the chromosome (Maynard Smith et al., 1991; Maiden et al., 1998; Feil et al., 1999).

Fourth, recombination in bacteria is not limited to the transfer of homologous segments. Bacteria can "capture" new gene loci from other organisms, sometimes from organisms that are extremely distantly related. This may occur as a side effect of homologous recombination, whereby a heterologous gene from a donor is integrated along with closely flanking homologous DNA (Hamilton et al., 1989; Majewski and Cohan, 1999a). Alternatively, heterologous genes may be integrated along with a transposable element introduced into the recipient cell on a plasmid or phage (Van Spanning et al., 1995).

I next consider the consequences of rare, promiscuous, localized recombination, both homologous and heterologous, on the properties of species and speciation in bacteria.

FORCES OF COHESION

Recombination has long been thought to be a powerful force of cohesion within a plant or animal species. Occasional recombination between freely interbreeding populations is thought to impose a cap on genetic divergence within a species (Mayr, 1982), although the mechanism by which this occurs is not clear (Futuyma, 1987). However, recombination is not the only force of cohesion acting on a plant or animal species. Members of a species are demographically interchangeable (sensu Templeton, 1989), so that through genetic drift and founder events any one individual can become the ancestor of the whole species in the future. Recently, sequence surveys of intraspecific variation have shown that drift and founder events contribute significantly to limiting diversity over the entire geographic range of an animal species (Avise and Walker, 1999; Avise, 2000).

What, then, are the forces of cohesion within a rarely sexual bacterial species with extremely large population sizes, where neither genetic exchange nor drift is likely to constrain diversity significantly? One could imagine that an asexual or rarely sexual species might have no cap on its divergence, such that a closely related group of bacteria would grow indefinitely in its diversity of sequences and phenotypes. However, this is clearly not the case. Bacterial clusters typically have a modest level of sequence diversity, usually an average sequence divergence of \sim 1% or less (Palys et al., 1997), not very different from that of many eukaryote species (Avise, 2000).

Diversity among a cluster of asexual (or rarely sexual) bacteria is capped by a force peculiar to rarely recombining organisms. In a rarely sexual population, an adaptive mutation sets in motion a round of natural selection that purges the population of nearly all its diversity at all loci. In the absence of recombination, the original adaptive mutant and its clonal descendants eventually replace the other cells of the population, with which they compete. Because the entire genome originally associated with the adaptive mutation remains intact as it sweeps through the population, the population loses its genetic diversity at all loci. Every adaptive mutation in an asexual population has the potential to set in motion this purging of diversity, a process known as "periodic selection" (Atwood et al., 1951) or "selective sweep" (Guttman and Dykhuizen, 1994). The recombination rates typical of bacteria are expected to allow purging of 99.9% of diversity at each locus (Cohan, 1994a).

In line with my earlier work, I will define a bacterial "ecotype" with respect to the fate of an adaptive mutant. (I previously referred to an "ecotype" as an "ecological population;" Cohan, 1994b.) An ecotype is a set of strains using the same or very similar ecological niches, such that an adaptive mutant from within the ecotype outcompetes to extinction all other strains of the same ecotype; an adaptive mutant does not, however, drive to extinction strains from other ecotypes.

Given periodic selection, a bacterial ecotype is not expected to expand its diversity indefinitely. The process of periodic selection should yield a cycling of diversity levels, whereby diversity is accumulated and then

a wave of periodic selection crashes the accumulated diversity back to near zero. Thus, divergence within a population of cells subjected to one another's adaptive mutations is only temporary, awaiting its demise with the next periodic selection event (Cohan, 1994a, 1994b, 1995, 1996, 1999).

What, then, is required for the creation of *permanent* divergence among closely related bacteria?

THE ORIGINS OF PERMANENT DIVERGENCE

An ecotype may expand its ecological diversity through accumulation of mutations and by receiving genes from other ecotypes. As long as the ecological differences among genotypes within the ecotype are only minor, a periodic selection event may purge the ecological and genetic diversity within the ecotype. Divergence will be permanent only if a new mutant (or recombinant) can escape the diversity-purging effects of periodic selection within its former ecotype. For example, if a recombinant acquires a new locus that allows metabolism of a new resource, the nascent and ancestral populations may come to overlap very little in their ecological niches. In this case, the new genotype and its descendants will form a separate ecotype, no longer susceptible to the periodic selection events within the parental ecotype. Each of the two ecotypes can now have its own private periodic selection events. At this point, periodic selection will purge the diversity only within and not between the ecotypes. Periodic selection will thereby keep the ecotypes distinct in all their characteristics and in all their genes (Fig. 1). The ecotypes are now free to diverge indefinitely, without the constraint of periodic selection.

What other forces might possibly constrain the divergence of nascent ecotypes? In highly sexual organisms, recombination between nascent species is a major obstacle to speciation. For two highly sexual eukaryotic species to coexist, the rate of between-population recombination must be severely reduced from the rate of recombination within populations. In contrast, recurrent recombination between nascent bacterial ecotypes is not expected to hinder their divergence greatly (Cohan, 1994b). Given the extremely low rate of recombination observed in bacteria, the influx of genes

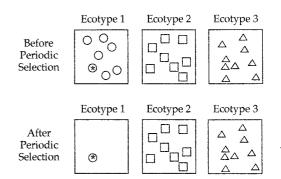


FIGURE 1. The transience of diversity within ecotypes, and the permanence of divergence among ecotypes. Each individual symbol represents an individual organism, and the distance between symbols represents the sequence divergence between organisms. The asterisk represents an adaptive mutation in one individual of Ecotype 1. Because of the rarity of recombination, the adaptive mutant and its clonal descendants replace the rest of the genetic diversity within the ecotype. Genetic diversity within an ecotype is thus transient, awaiting its demise with the next periodic selection event. Because ecotypes differ in the resources they use, the adaptive mutant from one ecotype (e.g., Ecotype 1) does not outcompete the cells of other ecotypes, and the genetic diversity within these other ecotypes remains untouched. Once populations of cells are divergent enough to escape one another's periodic selection events (i.e., the populations represent separate ecotypes), those populations are free to diverge permanently. Used with permission from the American Society for Microbiology (Cohan, 1996).

from other ecotypes cannot disrupt the integrity of each ecotype's specific adaptations (Cohan, 1994b). Even if recombination between ecotypes occurs at the same rate as recombination within ecotypes, selection against interecotypic recombinants can easily limit the frequency of recombinant genotypes to negligible values. The evolution of sexual isolation is not a necessary step in the origin of new ecotypes (Cohan, 1994b).

SPECIES AND SPECIATION

It appears, then, that the bacterial ecotypes described here share some fundamental dynamic properties of species in the eukaryotic world. Periodic selection imposes a cap on the phenotypic and sequence divergence within a bacterial ecotype, comparable to forces of cohesion acting within highly sexual eukaryotic species (i.e., genetic exchange and genetic drift). Bacterial ecotypes are thus groups of organisms the divergence of which is opposed by some force of cohesion, and so ecotypes are consistent with the species defined by the Cohesion Species Concept (Meglitsch, 1954; Templeton, 1989), as

well as by other modern species concepts (de Queiroz, 1998).

Of course, bacterial ecotypes are also consistent with the species of Van Valen's (1976) Ecological Species Concept, in which species are defined as groups holding different ecological niches.

As is the case for highly sexual eukaryotes, there is a threshold level of divergence beyond which two bacterial populations break free of cohesive forces and establish their separate evolutionary fates. Speciation in highly sexual eukaryotes requires a threshold level of reproductive as well as ecological divergence. For bacteria, the critical step in the origin of ecotypes is to break free of the periodic selection events occurring within the parental ecotype; at this point, nascent ecotypes are free to diverge without constraint (Cohan, 1994b). Bacterial ecotypes thus hold the quintessential property of all modern species concepts: that ecotypes are evolutionary lineages which are irreversibly separate, each with its own evolutionary tendencies and historical fate (Simpson, 1961; Wiley, 1978; de Queiroz, 1998). A species in the bacterial world may be understood as the evolutionary lineage held together by ecotypespecific periodic selection.

This is not the first suggestion to define bacterial species by their dynamic properties. Dykhuizen and Green (1991) earlier proposed to classify bacteria by the Biological Species Concept, delimiting bacterial species as groups of strains that recombine with one another but not with strains from other such groups, such that "phylogenies of different genes from individuals of the same species should be significantly different, whereas phylogenies of genes from individuals of different species should not be significantly different" (Dykhuizan and Green, page 7266, 1991). Indeed, it is often the case that phylogenetic evidence of recombination is much more common within named species than between them (Dykhuizen and Green, 1991). However, although bacteria can be classified into species according to patterns of recombination, there is no biological motivation for doing so. In contrast to the eukaryotes, genetic exchange between bacterial species does not hinder adaptive divergence, so the rise of sexual isolation is not an important milestone toward permanent divergence. Indeed, bacterial species that exchange genes are free to diverge without bound in all

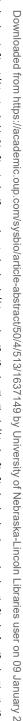
characters, neutral and adaptive (Cohan, 1994a, 1994b, 1995, 1996, 1999; Majewski and Cohan, 1999b).

Perhaps most importantly, the ecotype concept of bacterial species allows easy discovery of species, as well as classification of unknown strains into characterized species, even when we are ignorant of the ecology, as I describe next.

A CONVENIENT METHOD FOR CLASSIFYING AND DISCOVERING BACTERIAL ECOTYPES

Palys et al. (1997) predicted that, given enough time, bacterial ecotypes should be identifiable as sequence clusters, where the average sequence divergence between ecotypes is expected to be much greater than the average sequence divergence within them, for any gene shared among ecotypes; moreover, each ecotype should eventually be identifiable as a monophyletic group in a sequence-based phylogeny. The rationale can be outlined from a phylogenetic perspective. Suppose a new ecotype is derived clonally from one mutant cell that is adapted to a new ecological niche. Also suppose (for the time being) that these ecotypes are entirely asexual. The new ecotype is a monophyletic group descending from this original mutant, but this ecotype is not yet a sequence similarity cluster (Fig. 2a). This is because the average sequence divergence within the new ecotype is not much less than the average sequence divergence between the new ecotype and the most closely related clade from the ancestral ecotype. After the first periodic selection event, however, the diversity within the new ecotype is purged (Fig. 2b). Likewise, periodic selection events in the ancestral ecotype will purge diversity within that ecotype as well (Figs. 2c, 2d), and periodic selection in the ancestral ecotype renders that ecotype monophyletic. Thus, each ecotype will eventually appear as a monophyletic group and as a distinct cluster (Fig. 2e). Although this result is seen most clearly in the case of no recombination, Palys et al. (1997) showed that, under the extremely low rates of recombination occurring in bacteria, different ecotypes are still expected to fall eventually into different sequence clusters for any gene shared across ecotypes.

Conversely, ecotypes will not generally be split into two or more sequence clusters



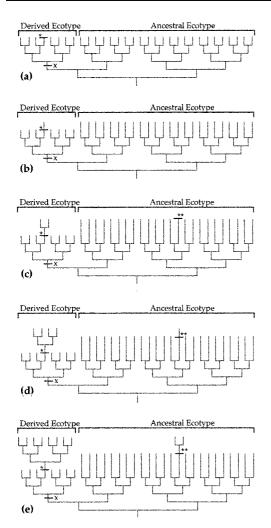


FIGURE 2. A phylogenetic perspective on periodic selection. As demonstrated here, two ecotypes will become distinct sequence similarity clusters, where the betweenecotype divergence is much greater than the withinecotype divergence. In these phylograms the vertical distance corresponds to evolutionary time and sequence divergence. (a) The derived ecotype consists of the descendants of a mutant (X) capable of utilizing a new ecological niche. The adaptive mutant in the derived ecotype (*) is capable of outcompeting all other members of the derived ecotype. Note that at this point the two ecotypes do not appear as distinct clusters. Moreover, the ancestral ecotype does not appear as a monophyletic group. (b) The adaptive mutant (*) has driven all the other lineages within the derived ecotype to extinction. (c) With time, the derived ecotype becomes more genetically diverse. One cell in the ancestral ecotype (**) has developed a mutation that allows it to outcompete other members of its ecotype. (d) The adaptive mutant (**) has outcompeted other members of the ancestral ecotype. (e) The ancestral ecotype is becoming more genetically diverse. At this point, each ecotype is a distinct sequence cluster as well as a monophyletic group.

(Palys et al., 1997; Dykhuizen, 1998), because multiple clusters within an ecotype would be unstable with respect to periodic selection. Each adaptive mutant within the ecotype would drive to extinction cells from all the clusters of the ecotype, such that the cluster bearing the adaptive mutant would be all that survives this purge of diversity. If two highly divergent clusters have coexisted long enough to survive periodic selection, then the clusters must belong to different ecotypes. On the other hand, two allopatric sequence clusters could belong to the same ecotype, because an adaptive mutant from one geographic region would have no opportunity to outcompete allopatric members of the same ecotype.

In summary, sequence clusters are expected to correspond, more or less, to ecotypes. This predicted correspondence between ecologically distinct groups and sequence clusters has been corroborated by a survey of sequence and ecological diversity in dozens of bacterial taxa (Palys et al., 1997).

The correspondence between ecologically distinct populations and sequence clusters has proved useful for bacterial systematics in two ways (Palys et al., 2000). First, established sequence differences between taxa can be used diagnostically to identify unknown isolates from the environment. For example, multilocus sequence typing (MLST) of protein-coding genes allows rapid diagnosis of pathogens by sequence data over the Internet (Maiden et al., 1998). Also, because nearly every bacterial species has been sequenced for 16S rRNA genes, any new isolate can readily be placed on the universal 16S rRNA-based tree (Goodfellow et al., 1997).

Second, the correspondence between ecotypes and sequence clusters is useful for discovering cryptic ecological diversity within a taxon. In several cases, a survey of sequence diversity within a named species has revealed multiple sequence clusters that later were found to be ecologically distinct (Palys et al., 1997). For example, sequence and allozyme surveys of *Borrelia burgdorferi* (sensu lato) led to the discovery of several Borrelia species with different pathogenic properties (Baranton et al., 1992; Balmelli and Piffaretti, 1996). Because protein-coding genes evolve more quickly than 16S rRNA, protein-coding genes may prove more useful than 16S rRNA

for discovering cryptic ecological diversity within named species (Palys et al., 2000).

Consider one additional challenge to discovering ecotypes through sequence clusters: Most named "species," both prokaryotic and eukaryotic, contain a hierarchy of sequence clusters and subclusters. This opens the possibility that a typical, named bacterial species may contain many cryptic and uncharacterized ecotypes, each corresponding to some small subcluster. Our task is to determine which level of subcluster, if any, actually corresponds to ecotypes. Fortunately, the peculiar population dynamics of bacterial ecotypes allows us to approach sequence-based discovery of ecotypes with greater confidence than is possible in highly sexual macroorganisms.

Jason Libsch and I have worked out a method for identifying the clusters that correspond to ecotypes (unpubl. results). Our "star phylogeny" approach assumes that the sequence diversity within an ecotype is limited nearly entirely by periodic selection and hardly at all by genetic drift. The rationale for this assumption is that, given the enormous sizes of bacterial populations, the average sequence divergence within an ecotype is much too small to be explained by genetic drift. Thus, nearly all strains randomly sampled from one ecotype should trace their ancestries back to the adaptive mutant that caused and survived the last periodic selection event. Thus, the phylogeny of an ecotype should be consistent with a star radiation, with all members of the ecotype equally closely related to one another.

We have developed a Monte Carlo simulation for determining whether a sequencebased phylogeny for a species is consistent with a star radiation, allowing for homoplasy by way of mutation and recombination. When a species' phylogeny shows too many internal nodes to be consistent with a star, we can conclude that multiple ecotypes must be present. Individual ecotypes can be tentatively identified as the most inclusive clades that are each consistent with the expectations for the simulation for one ecotype. Our preliminary analyses suggest the existence of numerous ecotypes within named bacterial species. This approach may provide an opportunity for bacterial systematists to characterize a much a greater wealth of ecological diversity than previously expected, even within named species.

How Many Bacterial Species Are There?

The International Committee on Systematic Bacteriology currently recognizes 5,236 bacterial species, each of which has been cultured and characterized (Euzéby, 1997 [updated 2000]). However, this accounting of bacterial diversity is limited to the bacteria that microbiologists know how to culture. Surveys of 16S rRNA variation from environmental samples suggest that culturable bacteria amount to <0.1% of the species in a given environment (Ward et al., 1998; Hugenholtz et al., 1998; Dojka et al., 2000; DeLong and Pace, this issue). Torsvik et al. (1998) assessed bacterial diversity by measuring the rates of annealing of heterogeneous DNA extracted from soil and found evidence for at least 10,000 bacterial species in 1 g of forest soil. Beginning with the results of Torsvik et al. (1998), Dykhuizen (1998) has argued that worldwide there must be at least 109 bacterial species. We may further plausibly argue that each of the world's animal and plant species has at least one species-specific pathogen. For example, our own species has several bacterial pathogens infecting only humans, including Streptococcus pyogenes and some related Streptococci (Davis et al., 1990). If *Homo sapiens* is not unusual, and most eukaryotic species bear their own species-specific pathogens, the number of bacterial species would be large indeed.

POTENTIAL FOR SPECIATION IN THE BACTERIAL WORLD

A large number of bacterial species may be explained by an enormous potential for speciation in the bacterial world. Many attributes of bacterial genetic exchange, plus the large population sizes typically found in bacteria, should allow much greater opportunity for speciation than is possible in the highly sexual animals and plants.

First, speciation in highly sexual eukaryotes requires both reproductive (Mayr, 1982) and ecological (Gause, 1934; Eldredge, 1985) divergence, but speciation in bacteria requires only ecological divergence (Cohan, 1994b).

Second, the extremely large population sizes of bacteria make rare mutation and recombination events much more accessible to a bacterial population than is the case for macroorganisms.

Third, whereas plant and animal species are genetically closed to all other species (or at least closed to all but the most closely related species with which hybridization is possible), a bacterial species is open to gene transfers from many taxa, even those distantly related (Young, 1989). So, while plant and animal species must evolve all their adaptations on their own, bacteria can take up existing adaptations from anywhere in the bacterial world. In some cases, homologous recombination can substitute an adaptive allele from another species into an existing gene in the recipient. For example, a short segment of an allele for penicillin resistance has been transferred by homologous recombination from Neisseria flavescens into N. gonorrheae (Maynard Smith et al., 1991). Recombination can also introduce entirely novel genes and operons from other species (Lan and Reeves, 1996; Groisman and Ochman, 1997; Lawrence and Ochman, 1998; Alm et al., 1999). This ready acquisition of new adaptations should make invasion of new niches easier for bacteria than eukaryotes.

The transfer of adaptations across species is facilitated by several aspects of bacterial genetic exchange. Integration of highly divergent DNA is made possible first by the promiscuous nature of bacterial genetic exchange. Also important is the localized nature of recombination events, whereby only a very small fraction of the donor's genome is integrated. This allows for transfer of a generally useful adaptation (i.e., useful in the genetic backgrounds and the ecological niches of both the donor and the recipient), without the cotransfer of narrowly adaptive donor segments that would be deleterious for the recipient (Zawadzki and Cohan, 1995). This is in contrast to the case for eukaryotes, where the processes of meiosis and fertilization yield hybrids that are a 1:1 mix of both parents' genomes. Also, perhaps the physiology of bacteria is more modular than is the case for macroorganisms, in that an entirely new adaptation might be accommodated with very little fitness cost.

Fourth, the possibility of heterologous transfer greatly facilitates the escape of a nascent ecotype from competition with its former population. By taking on an entirely new metabolic function, a nascent ecotype is likely to utilize resources not available to its parental ecotype. This would instanta-

neously put the nascent ecotype out of range of periodic selection events in its former ecotype.

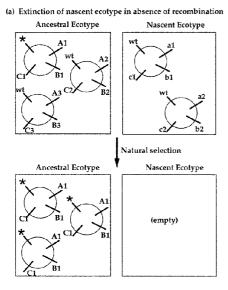
Finally, the promiscuity of bacterial genetic exchange may help a nascent ecotype from being extinguished by an adaptive mutant from the parental ecotype. Let us suppose that the ecological divergence of a nascent ecotype from the parental ecotype requires several evolutionary steps. In the early stages of such divergence, most periodic selection events may be limited to each ecotype, but occasionally an extraordinarily competitive adaptive mutant from the parental ecotype might outcompete all strains from the nascent ecotype (as well as strains from its own ecotype). In this case, genetic exchange between the ecotypes could prevent the extinction of one ecotype by the other (Fig. 3): Through genetic exchange, the nascent ecotype would acquire the adaptive mutation from the other ecotype, and the new ecotype would thereby lose its disadvantage.

Note that although recombination in bacteria is not frequent enough to prevent ecological divergence (Cohan, 1994b) or neutral sequence divergence (Cohan, 1995; Palys et al., 1997) between ecotypes, it should be frequent enough to allow an adaptive mutation to pass between ecotypes and initiate a periodic selection event in the recipient ecotype. Prevention of divergence (whether neutral or adaptive) among ecotypes requires recurrent recombination, and bacterial rates of recombination are not up to this task (Cohan, 1994b). However, initiating a selective sweep requires only a *single* recombinational transfer of the adaptive mutation into the recipient ecotype; so, given the enormous population sizes of bacterial populations, such a transfer event is not unlikely.

Ironically, whereas recombination between nascent, highly sexual species hinders their speciation, the low level of recombination between nascent bacterial species might promote their coexistence and allow further divergence.

CONCLUSIONS AND FUTURE DIRECTIONS

Defining bacterial species as "ecotypes" (i.e., according to the domain of extinction caused by periodic selection events) has several attractive properties. Such species would be ecologically distinct from one another; each species would be subject to a



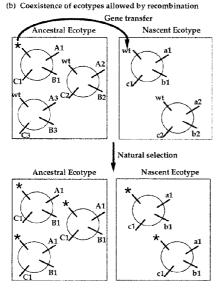


FIGURE 3. Facilitation of speciation by recombination among ecotypes. It is assumed that each newly divergent ecotype in the figure has already undergone several private periodic selection events. However, in the figure we suppose that an extraordinarily competitive adaptive mutant (with mutation) has appeared in the ancestral ecotype, such that this mutant would outcompete the membership of the nascent ecotype as well as its own ecotype membership. (a) When there is no recombination between the newly divergent ecotypes, the adaptive mutant could extinguish the membership of the other ecotype, and the speciation process would be terminated. (b) When the adaptive mutation can be transferred from one ecotype to the other, periodic selection is less likely to cause extinction of one ecotype by another. The transfer of the adaptive mutation would cause a private periodic selection event within the nascent ecotype. Because the two populations would then share the adaptive mutation, one ecotype would not be able to extinguish the other.

powerful force of cohesion (periodic selection); and different species would be free to diverge without constraint from one another. Moreover, the correlation between ecotypes and sequence clusters allows easy discovery and identification of ecotypes even when we are ignorant of their ecology.

Speciation in bacteria appears much less constrained than speciation in the eukaryotic world, and so we might expect rates of speciation to be much greater for bacteria than for eukaryotes. Little is known about extinction in the bacterial world, although we might expect it to be rarer for bacterial species. The much larger population sizes of bacteria at least prevent extinctions attributable to demographic stochasticity (Dykhuizen, 1998). We should not be surprised to discover and characterize, eventually, a much greater number of bacterial species than the 5,236 currently valid species. Recent discoveries of taxa by culture-free, sequence-based methods (Hugenholtz et al., 1998; Dojka et al., 2000) and in situ approaches toward characterizing the ecology of unculturable taxa (Ramsing et al., 2000) are promising steps in this direction.

Speciation in the bacterial world appears to be limited by mutational and recombinational events that allow invasions into new niches. Exploring the rate of occurrence of such mutations, in comparison with the rate of mutations causing periodic selection within an ecotype, will be interesting. Because bacterial speciation is also limited by the extent to which nascent ecotypes are able to escape all periodic selection events stemming from their former ecotype, it will be interesting to explore the role of recombination in facilitating this escape process. These issues can be addressed using recently developed experimental systems for studying bacterial speciation (Rainey and Travisano, 1998; Treves et al., 1998).

It is not yet clear which is more important in bacterial niche invasion—mutations in existing genes or acquisition of new gene loci from other species. If we should find that a typical named species contains only one ecotype, gene acquisition probably plays the major role. Because closely related species frequently bear very different sets of gene loci (Lan and Reeves, 1996; Groisman and Ochman, 1997; Lawrence and Ochman, 1998; Alm et al., 1999), these differences appear to

contribute significantly to ecological divergence between the species (Groisman and Ochman, 1997). On the other hand, we may find evidence of many ecotypes within a typical named species. In this case, the most closely related ecotypes would have very few differences in their complements of gene loci, and the genetic basis of ecological differentiation between such ecotypes might then involve only changes in existing gene loci.

ACKNOWLEDGMENTS

I am grateful to Kevin de Queiroz and Daniel Dykhuizen for critically reading the manuscript and for making many helpful suggestions. This work was supported by U.S. Environmental Protection Agency grant R82-5348-01-2.

REFERENCES

- ALM, R. A., L. L. LING, D. T. MOIR, B. L. KING, E. D. BROWN, P. C. DOIG, D. R. SMITH, B. NOONAN, B. C. GUILD, B. L. DEJONGE, ET AL. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature 397:176–180
- ATWOOD, K. C., L. K. SCHNEIDER, AND F. J. RYAN. 1951. Periodic selection in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 37:146–155.
- AVISE, J. C. 2000. Phylogeography: The history and formation of species. Harvard Univ. Press Cambridge, Massachusetts.
- AVISE J. C., AND D. WALKER. 1999. Species realities and numbers in sexual vertebrates: Perspectives from an asexually transmitted genome. Proc. Natl. Acad. Sci. USA 96:992–995.
- BALMELLI, T., AND J. PIFFARETTI. 1996. Analysis of the genetic polymorphism of *Borrelia burgdorferi* sensu lato by multilocus enzyme electrophoresis. Int. J. Syst. Bacteriol. 46:167–172.
- BARANTON, G., D. POSTIC, I. SAINT GIRONS, P. BOERLIN, J. C. PIFFARETTI, M. ASSOUS, AND P. A. GRIMONT. 1992. Delineation of *Borrelia burgdorferi* sensu stricto, *Borrelia garinii* sp. nov., and group VS461 associated with Lyme borreliosis. Int. J. Syst. Bacteriol. 42:378–383.
- BARRETT, S. J., AND P. H. A. SNEATH. 1994. A numerical phenotypic taxonomic study of the genus *Neisseria*. Microbiology 140:2867–2691.
- BRIDGE, P. D., AND P. H. A. SNEATH. 1983. Numerical taxonomy of *Streptococcus*. J. Gen. Microbiol. 129:565–597.
- BROCK, T. D. 1978. Thermophilic microorganisms and life at high temperatures Springer-Verlag, New York.
- BROOM, A., AND P. H. SNEATH. 1981. Numerical taxonomy of *Haemophilus*. J. Gen. Microbiol. 126:123–149.
- COHAN, F. M. 1994a. Genetic exchange and evolutionary divergence in prokaryotes. Trends Ecol. Evol. 9:175– 180.
- COHAN, F. M. 1994b. The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. Am. Nat. 143:965–986.

- COHAN, F. M. 1995. Does recombination constrain neutral divergence among bacterial taxa? Evolution 49:164–175.
- COHAN, F. M. 1996. The role of genetic exchange in bacterial evolution. Am. Soc. Microbiol. News 62:631–636
- COHAN, F. M. 1999. Genetic structure of bacterial populations. Pages 475–489 *in* Evolutionary genetics from molecules to morphology (R. Singh and C. Krimbas, eds.). Cambridge Univ. Press, Cambridge.
- COHAN, F. M., M. S. ROBERTS, AND E. C. KING. 1991. The potential for genetic exchange by transformation within a natural population of *Bacillus subtilis*. Evolution 45:1393–1421.
- DAVIS, B. D., R. DULBECCO, H. N. EISEN, AND H. S. GINSBERG. 1990 Microbiology, 4th edition, J. B. Lippincott, Philadelphia.
- DE QUEIROZ, K. 1998. The general lineage concept of species, species criteria, and the process of speciation. Pages 57–75 *in* Endless forms: Speciation and speciation (D. J. Howard and S. H. Berlocher, eds.). Oxford Univ. Press, Oxford.
- DOJKA, M. A., J. K. HARRIS, AND N. R. PACE. 2000. Expanding the known diversity and environmental distribution of an uncultured phylogenetic division of Bacteria. Appl. Environ. Microbiol. 66:1617–1621.
- DUNCAN, K. E., C. A. ISTOCK, J. B. GRAHAM, AND N. FERGUSON. 1989. Genetic exchange between *Bacillus subtilis* and *Bacillus licheniformis*: Variable hybrid stability and the nature of species. Evolution 43:1585–1609.
- DYKHUIZEN, D. E. 1998. Santa Rosalia revisited: Why are there so many species of bacteria? Antonie Van Leeuwenhoek 73:25–33.
- DYKHUIZEN, D. E., AND L. Green. 1991. Recombination in *Escherichia coli* and the definition of biological species. J. Bacteriol. 173:7257-7268.
- EDWARDS, R. A., R. A. HELM, AND S. R. MALOY. 1999. Increasing DNA transfer efficiency by temporary inactivation of host restriction. BioTechniques 26:892–898.
- ELDREDGE, N. 1985. Unfinished synthesis: Biological hierarchies and modern evolutionary thought. Oxford Univ. Press, New York.
- EMBLEY, T. M., AND E. STACKEBRANDT. 1997. Species in practice: Exploring uncultured prokaryote diversity in natural samples. Pages 61–81 *in* Species: The units of biodiversity (M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds.). Chapman & Hall, London.
- EUZÉBY, J. P. 1997. List of bacterial names with standing in nomenclature: A folder available on the Internet (last updated: July 2000, URL: http://www.sv.cict.fr/bacterio/). Int. J. Syst. Bacteriol. 47:590–592.
- FEIL, E. J., M. C. 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. 16:1496–502.
- FENCHEL, T. AND B. J. FINLAY. 1995. Ecology and evolution in anoxic worlds. Oxford Univ. Press, Oxford.
- FUTUYMA, D. J. 1987. On the role of species in anagenesis. Am. Nat. 130:465–473.
- GAUSE, G. F. 1934. The struggle for existence. Williams & Wilkins, Baltimore.
- GOODFELLOW, M., G. P. MANFIO, AND J. CHUN. 1997. Towards a practical species concept for cultivable bacteria. Pages 25–59 *in* Species: The units of biodiversity. (M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds.). Chapman & Hall, London.

- GROISMAN, E. A., AND H. OCHMAN. 1997. How *Salmonella* became a pathogen. Trends Microbiol. 5:343–349.
- GUTTMAN, D. S., AND D. E. DYKHUIZEN. 1994. Detecting selective sweeps in naturally occurring *Escherichia coli*. Genetics 138:993–1003.
- HAMILTON, C. M., M. ALDEA, B. K. WASHBURN, P. BABITZKE, AND S. R. KUSHNER. 1989. New method for generating deletions and gene replacements in Escherichia coli. J. Bacteriol. 171:4617–4622.
- HEY, J., AND J. WAKELEY. 1997. A coalescent estimator of the population recombination rate. Genetics 145:833– 846
- HSIEH, P., C. S. CAMERINI-OTERO, AND R. D. CAMERINI-OTERO. 1992. The synapsis event in the homologous pairing of DNAs: RecA recognizes and pairs less than one helical repeat of DNA. Proc. Natl. Acad. Sci. USA 89:6492–6496.
- HUDSON, R. R. 1987. Estimating the recombination parameter of a finite population model without selection. Genet. Res. 50:245–250.
- HUDSON, R. R., AND N. L. KAPLAN. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111:147–164.
- HUGENHOLTZ, P., B. M. GOEBEL, AND N. R. PACE. 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J. Bacteriol. 180:4765–4774.
- JOHNSON, J. L. 1973. Use of nucleic-acid homologies in the taxonomy of anaerobic bacteria. Int. J. Syst. bacteriol. 23:308–315.
- LAN, R., AND P. R. REEVES. 1996. Gene transfer is a major factor in bacterial evolution. Mol. Biol. Evol. 13:47–55.
- LAWRENCE, J. G., AND H. OCHMAN. 1998. Molecular archaeology of the *Escherichia coli* genome. Proc. Natl. Acad. Sci. USA 95:9413–9417.
- MADIGAN, M. T., AND B. L. MARRS. 1997. Extremophiles. Sci. Am. 276:82–87.
- MADIGAN, M. T., J. M. MARTINKO, AND J. PARKER. 1999. Brock's biology of microorganisms, 9th edition, Prentice-Hall Inc., Upper Saddle River, NJ.
- MAIDEN, M. C., J. A. BYGRAVES, E. FEIL, G. MORELLI, J. E. RUSSELL, R. URWIN, O. ZHANG, J. ZHOU, K. ZURTH, D. A. CAUGANT, ET AL. 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. USA 95:3140–3145.
- 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. M. COHAN. 1999a. DNA sequence similarity requirements for interspecific recombination in Bacillus. Genetics 153:1525–1533.
- MAJEWSKI, J., AND F. M. COHAN. 1999b. Adapt globally, act locally: The effect of selective sweeps on bacterial sequence diversity. Genetics 152:1459–1474.
- MAJEWSKI, J., P. ZAWADZKI, P. PICKERILL, F. M. COHAN, AND C. W. DOWSON. 2000. Barriers to genetic exchange between bacterial species: Streptococcus pneumoniae transformation. J. Bacteriol. 182:1016–1023.
- MALLET, J. 1995. A species definition for the modern synthesis. Trends Ecol. Evol. 10:294–299.
- MARGULIS, L., AND K. V. SCHWARTZ. 1998. Five kingdoms: An illustrated guide to the phyla of life on earth. Freeman, New York.
- MAY, R. M. 1973. On relationships among various types of population models. Am. Nat. 107:46–57.

- MAYNARD SMITH, J., C. G. DOWSON, AND B. G. SPRATT. 1991. Localized sex in bacteria. Nature 349:29–31.
- 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.
- MAYR, E. 1982. The growth of biological thought: Diversity, evolution, and inheritance. Harvard Univ. Press, Cambridge, Massachusetts.
- MCKANE, M., AND R. MILKMAN. 1995. Transduction, restriction and recombination patterns in *Escherichia coli*. Genetics 139:35–43.
- MEGLITSCH, P. A. 1954. On the nature of species. Syst. Zool. 3:491–503.
- MILKMAN R., E. A. RALEIGH, M. MCKANE, D. CRYDERMAN, P. BILODEAU, AND K. MCWEENY. 1999. Molecular evolution of the *Escherichia coli* chromosome. V. Recombination patterns among strains of diverse origin. Genetics 153:539–54.
- MORITZ, C., C. J. SCHNEIDER, AND D. B. WAKE. 1992. Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. Syst. Biol. 41:273–291.
- Palys, T., É. Berger, I. Mitrica, L. K. Nakamura, and F. M. Cohan. 2000. Protein-coding genes as molecular markers for ecologically distinct populations: The case of two *Bacillus* species. Int. J. Syst. Evol. Microbiol. 50:1021–1028.
- PALYS, T., F. M. COHAN, AND L. K. NAKAMURA. 1997. Discovery and classification of ecological diversity in the bacterial world: The role of DNA sequence data. Int. J. Syst. Bacteriol. 47:1145–1156.
- RAINEY, P. B., AND M. TRAVISANO 1998. Adaptive radiation in a heterogeneous environment. Nature 394:69–72.
- RAMSING, N. B., M. J. FERRIS, AND D. M. WARD. 2000. Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter-thick photic zone of a hot spring cyanobacterial mat. Appl. Environ. Microbiol. 66:1038–1049.
- RAO, B. J., S. K. CHIU, L. R. BAZEMORE, G. REDDY, AND C. M. RADDING. 1995. How specific is the first recognition step of homologous recombination? Trends Biochem. Sci. 20:102–113.
- 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.
- ROBERTS, M. S., AND F. M. COHAN. 1993. The effect of DNA sequence divergence on sexual isolation in Bacillus. Genetics 134:401–408.
- ROBERTS, M. S., AND F. M. COHAN. 1995. Recombination and migration rates in natural populations of *Bacillus subtilis* and *Bacillus mojavensis*. Evolution 49:1081–1094.
- SELANDER, R. K., AND J. M. MUSSER. 1990. Population genetics of bacterial pathogenesis. Pages 11–36 *in* Molecular basis of bacterial pathogenesis. (B. H. Iglewski and V. L. Clark, eds.). Academic Press, San Diego, California.
- SIMPSON, G. G. 1961. Principles of animal taxonomy. Columbia Univ. Press, New York.
- SNEATH, P. H. A. 1977. The maintenance of large numbers of strains of microorganisms and the implications for culture collections. FEMS Microbiol. Lett. 1:333–334.
- SNEATH, P. H., AND M. STEVENS, 1985. A numerical taxonomic study of *Actinobacillus*, *Pasteurella* and *Yersinia*. J. Gen. Microbiol. 131:2711–2738.

- STACKEBRANDT, E., AND B. M. GOEBEL. 1994. Taxonomic note: A place for DNA:DNA reassociation and 16s rRNA sequence analysis in the present species definition in bacteriology. Int. J. Syst. Bacteriol. 44:846–849.
- TEMPLETON, A. R. 1989. The meaning of species and speciation: A genetic perspective. Pages 3–27 *in* Speciation and its consequences (D. Otte and J. A. Endler, eds.). Sinauer, Sunderland, Massachusetts.
- TORSVIK, V., F. L. DAAE, R. A. SANDAA, AND L. OVREAS. 1998. Novel techniques for analysing microbial diversity in natural and perturbed environments. J. Biotechnol. 64:53–62.
- TRAUTNER, T. A., B. PAWLEK, S. BRON, AND C. ANAGNOSTOPOULOS. 1974. Restriction and modification in *B. subtilis*. Mol. Gen. Genet. 131:181–191.
- Treves, D. S., S. Manning, and J. Adams. 1998. Repeated evolution of an acetate cross feeding polymorphism in long-term populations of *Escherichia coli*. Mol. Biol. Evol. 15:789–797.
- VAN SPANNING, R. J., W. N. REIJNDERS, AND A. H. STOUTHAMER. 1995. Integration of heterologous DNA into the genome of *Paracoccus denitrificans* is mediated by a family of IS1248-related elements and a second type of integrative recombination event. J. Bacteriol. 177:4772–4778.
- VAN VALEN, L. 1976. Ecological species, multispecies, and oaks. Taxon 25:233–239.
- VULIC, M., F. DIONISIO, F. TADDEI, AND M. RADMAN. 1997. Molecular keys to speciation: DNA polymor-

- phism and the control of genetic exchange in enterobacteria. Proc. Natl. Acad. Sci. USA 94:9763–9767.
- WARD, D. M., M. J. FERRIS, S. C. NOLD, AND M. M. BATESON. 1998. A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. Microbiol. Mol. Biol. Rev. 62:1353–1370.
- WAYNE, L. G., D. J. BRENNER, R. R. COLWELL, P. A. D. GRIMONT, O. KANDLER, M. I. KRICHEVSKY, L. H. MOORE, W. E. C. MOORE, R. G. E. MURRAY, E. STACKEBRANDT, ET AL. 1987. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Int. J. Syst. Bacteriol. 37:463–464.
- WHITTAM, T. S., AND Ś. E. AKE. 1993. Genetic polymorphisms and recombination in natural populations of *Escherichia coli*. Pages 223–245 *in* Molecular paleopopulation biology. (N. Takahata and A. G. Clark, eds.). Japan Scientific Society Press, Tokyo.
- WILEY, E. O. 1978. The evolutionary species concept reconsidered. Syst. Zool. 27:17–26.
- YOUNG, J. P. W. 1989. The population genetics of bacteria. Pages 417–438 *in* Genetics of bacterial diversity (D. A. Hopwood and K. F. Chater, eds.). Academic Press, London.
- ZAWADZKI, P., AND F. M. COHAN. 1995. The size and continuity of DNA segments integrated in Bacillus transformation. Genetics 141:1231–1243.
- ZAWADZKI, 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.

Received 14 April 2000; accepted 19 June 2000 Associate Editor: M. Kane