Current Biology *18*, R1024–R1034, November 11, 2008 DOI 10.1016/j.cub.2008.09.014

The Origins of Ecological Diversity in Prokaryotes

Review

Frederick M. Cohan and Alexander F. Koeppel

The urkingdoms and major divisions of prokaryotes are enormously diverse in their metabolic capabilities and membrane architectures. These ancient differences likely have a strong influence on the kinds of ecological adaptations that may evolve today. Some ecological transitions have been identified as having occurred primarily in the distant past, including transitions between saline and non-saline habitats. At the microevolutionary level, the likely existence of a billion prokaryotic species challenges microbiologists to determine what might promote rapid speciation in prokaryotes, and to identify the ecological dimensions upon which new species diverge and by which they may coexist. Rapid speciation in prokaryotes is fostered by several unique properties of prokaryotic genetic exchange, including their propensity to acquire novel gene loci by horizontal genetic transfer, as well as the rarity of their genetic exchange, which allows speciation by ecological divergence alone, without a requirement for sexual isolation. The ecological dimensions of prokaryotic speciation may be identified by comparing the ecology of the most newly divergent, ecologically distinct populations (ecotypes). This program is challenged by our ignorance of the physiological and ecological features most likely responsible for adaptive divergence between closely related ecotypes in any given clade. This effort will require development of universal approaches to hypothesize demarcations of ecotypes, and to confirm and characterize their ecological distinctness, without prior knowledge of a given clade's ecology.

Introduction

The discovery of the animal phylum Cycliophora in 1995 was deemed the 'zoological highlight of the decade' [1]. This unique addition to the stable of zoology was all the more remarkable for the ordinariness and familiarity of its habitat — the mouth of a lobster. That Cycliophora was so long overlooked raises the possibility that other animal phyla remain to be discovered. Nevertheless, it is the realm of the prokaryotes that holds the greatest promise for yielding previously unknown, profoundly divergent groups. Over the last two decades, microbial ecologists and systematists have searched familiar habitats, some as ordinary as sea water and forest soil, as well as other more exotic habitats, such as hot springs and deep sea vents, to discover the most anciently divergent of prokaryotes. The pace of discovery of deeply divergent prokaryotic groups has been feverish, and shows no signs of slowing in the near future (J. Tiedje, personal communication).

We might wonder why zoologists took so long to discover Cycliophora, but there is no mystery why major prokaryotic groups have been refractory to discovery: only a minute fraction of prokaryotes are at present cultivable [2], hindering laboratory study of physiology, and morphology is not a reliable indicator of prokaryotic diversity. Ecologists and systematists have worked around these difficulties through molecular surveys, usually by discovering taxa as distinct sequence clusters for the universal 16S rRNA gene, with or without cultivating the organisms [2].

The most astounding of such sequence-based discoveries was that all of cellular life fits into three 'urkingdoms', with the prokaryotic groups bacteria and archaea comprising two of these [3]. Systematists have become accustomed to the nearly quotidian discovery of bacterial 'divisions'. Divisions are the largest taxa within the bacteria, such as the cyanobacteria (the oxygen-producing photosynthetic bacteria), the spirochetes (corkscrew-shaped bacteria, including the pathogens that cause syphilis and Lyme disease), and the firmicutes (including Gram-positive bacteria), as well as many newly discovered, but uncultivated, 'candidate' divisions known only by a DNA sequence and a photograph. The depth of divergence among bacterial divisions is much greater than that among animal phyla, with the evolutionary split among most divisions pre-dating the origin of animals and many pre-dating the origin of the eukaryotes [2].

The prokaryotes appear to dwarf the eukaryotes in the number of species as well. Estimates of total eukaryotic diversity fall within the range of 10–50 million species [4]. Although only about 9000 species of prokaryotes have been described [5], indirect molecular approaches based on annealing of DNA extracted from the environment (without cultivation) suggest the existence of a billion or more prokaryotic species worldwide [6], and ten million species within a given habitat [7].

Here, we address the challenges that the enormous levels of prokaryotic diversity present to microbial ecologists and evolutionary biologists, from the origins of anciently divergent major groups to the everyday origins of prokaryotic species. We explain how the evolutionary changes leading to the major prokaryotic taxa are different from the changes associated with speciation today, and review methods for identifying the kinds of evolutionary changes that have occurred primarily in the distant past.

We present a microevolutionary perspective to studying the origins of prokaryotic species. It is with regard to speciation that we justify our focus on the prokaryotes, which are a paraphyletic group: the constituent bacteria and archaea share certain primitive characteristics of genetic exchange that are critical to their speciation. We review how the most closely related, ecologically distinct populations may be identified, and how their identification has shed light on the ecological dimensions on which prokaryotes diverge on speciation. We discuss why prokaryotes are expected to have higher rates of speciation than animals and plants, and why certain prokaryotic groups may speciate more frequently than others. Finally, we discuss how we can leverage what we already know, and can easily find out, to better characterize the origins of prokaryotic ecological diversity.

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

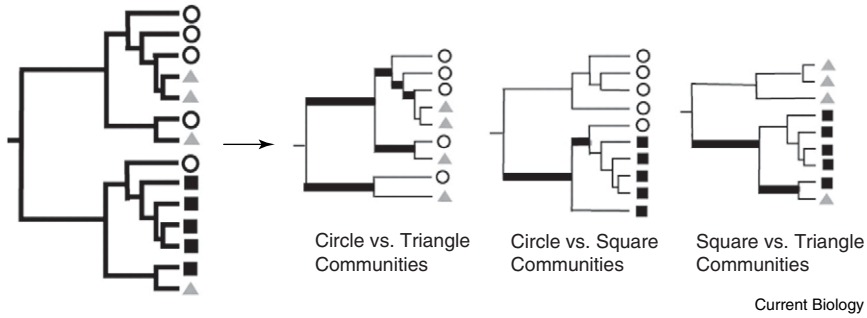

Figure 1. A phylogenetically based analysis of community differences.

The UniFrac metric quantifies community differences as the fraction of the phylogeny that is unique to each community. Each pair of communities is analyzed separately. Portions of each pairwise phylogeny that are *not* unique to a community are represented by wide bars. The members of three communities are indicated by different symbols. A high value of the UniFrac metric indicates that an evolutionary transition across communities is rare and ancient (for example, circles *versus* squares). Unifrac has the potential to help us discover the most difficult evolutionary changes, by focusing on the rarest and most ancient of transitions across habitat types. The UniFrac approach is limited at the moment to extremely broad habitat categories due to the imprecise nature of published habitat data [22]. The UniFrac analysis need not be limited to community differences; UniFrac could similarly be used to measure the difficulty of any kind of physiological or genomic transition in evolution. Adapted with permission from Lozupone and Knight [21].

Urkingdoms and Divisions

Mega-evolution of Major Taxa

The urkingdoms and divisions within urkingdoms are not just old; they are profoundly different in their metabolic capabilities. For example, while all photosynthetic eukaryotes use the oxygenic pathway (which was acquired from the cyanobacteria), the known photosynthetic divisions of prokaryotes utilize three fundamentally different pathways, with only the cyanobacteria yielding oxygen [8]. Also, while the mitochondrion-bearing eukaryotes use only oxygen as their electron acceptor — in the oxidative phosphorylation pathway, which was acquired from aerobic bacteria — various anaerobic bacteria use nitrate, sulfate, ferric, or aluminum (III) ions as electron acceptors [8], while some can even use organic molecules, such as trinitrotoluene [9]. These metabolic differences provide ecological opportunities for living in anoxic environments containing different oxidizing agents.

In the case of animals, the origins of the major taxa have involved ancient and unique, 'mega-evolutionary' reorganizations of fundamental developmental plans (or 'bauplans') [10]; likewise, the origins of major groups of prokaryotes appear to have involved ancient and unique changes in their bauplan-like structures, in this case the cell membranes. Indeed, the chemical fossil record shows that evolution of membrane lipid structure has been extremely conservative in prokaryotes [11,12], and the most radical transitions in cell membranes, such as the diester-diethyl transition between the archaea and bacteria, as well as the transition to a resilient cell wall found in the firmicutes, for example, were ancient and unique evolutionary events. We should not expect that microevolutionary studies of recent species differences will explain the ancient, mega-evolutionary origins of major groups, in either prokaryotes or animals [13].

Why would certain evolutionary transitions, such as the inventions of membrane bauplans, be extremely rare and difficult? The evolutionary origins of complex adaptations have been a challenge for evolutionary biology since Darwin [14], in that the origin of each individual component of a complex adaptation must yield an incremental advantage [15]. In prokaryotes, evolutionary origins are not as constrained, because horizontal genetic transfer can simultaneously bring an entire biochemical pathway from a donor genome into a recipient [16]. Indeed, horizontal gene transfer was partly responsible for the origins of complex features that separate bacterial divisions, such as flagellae [17]. Nevertheless, that certain ancient adaptations have never, or only rarely, been transferred — for example, 'signature' membrane differences between divisions [13] — suggests that some adaptations are either too complex to fit into a transferable segment, or perhaps that they are too tightly integrated into the rest of the cell's physiology and structure to be of use after transfer.

In both animals [18] and prokaryotes, ancient differences in structures can strongly influence the kinds of ecological adaptation that may evolve by microevolution in the here and now. For example, the glycerol diether bonds of archaeal membranes provide a more ready access to survivability at very high temperatures (>85°C) than is the case for the diester-bonded bacterial membranes; this is because the ether-linked lipids of archaea can produce a monolayer membrane that does not peal apart at very high temperatures [19].

Bauplan differences among divisions within the bacteria also offer different ecological opportunities. The resilient cell wall of the firmicutes confers tolerance to the osmotic stress of rapid re-wetting after drought. This does not rule out drought resistance in other bacterial groups, but their less resilient cell walls require fastidious attention to osmotic balance through fine tuning of osmolyte concentrations [20].

The Most Ancient Ecological Transitions

What are the ecological transitions that occurred primarily in the distant past? Our failure to cultivate most divisions limits our ability to predict their ecology, and to identify the most ancient and difficult ecological transitions (for example, those that occur only with the invention of a new division). Fortunately, two new approaches allow us to identify the most ancient ecological transitions without the benefit of either cultivation or study of physiology.

The phylogenetically based UniFrac approach can reveal the most ancient and difficult evolutionary changes by identifying the rarest and earliest of evolutionary transitions across habitat types (Figure 1) [21]. Lozupone and Knight [21] discovered that the transitions between saline and any of various non-saline habitats were the most ancient and infrequent; other ancient transitions were between aqueous and sediment habitats and between soil and sediment [22].

One may also identify ancient and difficult transitions by determining the environmental parameters that closely predict the relative abundances of higher taxon levels (for example, bacterial divisions). For example, Fierer *et al.* [23] found that carbon availability best predicted abundance of different divisions in soils, with Betaproteobacteria and Bacteriodetes being most common in soils with high carbon availability, and Acidobacteria most common in soils with

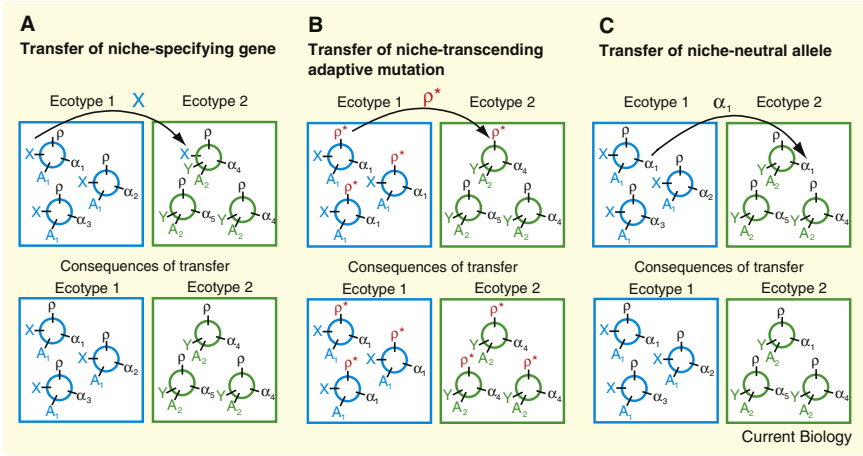

Figure 2. The consequences of recombination for genes of different fitness effects.

Ecotype 1 (blue) and Ecotype 2 (green) represent a pair of closely related, but ecologically distinct populations. Each contains a niche-specifying allele of gene *A*, with alleles A_1 and A_2 conferring niche-specifying adaptations in Ecotypes 1 and 2, respectively. Each ecotype also bears a unique niche-specifying locus, *X* or *Y*, in Ecotypes 1 and 2, respectively. Also shown are a niche-neutral locus α, as well as another locus ρ, at which a niche-transcending adaptive mutation, ρ^*, has occurred (in panel B). (A) The transfer of a niche-specifying gene between ecotypes does not disrupt the integrity of adaptive divergence between ecotypes. The equilibrium frequency of a niche-specifying gene from another ecotype is c_b/s, where c_b = the rate of recombination between ecotypes ($\sim 10^{-8}$ within ecotypes), and s = selective disadvantage of the recombinant [28]. Even without sexual isolation between ecotypes, the equilibrium frequency of recombinant niche-specifying genes from other ecotypes should be negligible. (B) When a niche-transcending adaptive mutation (ρ^*) is transferred between ecotypes, the result is a periodic selection event within the recipient ecotype. Ecotype 1 has recently undergone a selective sweep due to the adaptive mutation in the ρ locus (indicated by homogeneity at the α locus). The adaptive change (represented by an asterisk and the color red) is niche-transcending, in that the change is adaptive in either of the two niches. When ρ^* is then transferred into Ecotype 2, the result is a selective sweep in that ecotype, eliminating diversity at all loci. Even a single transfer of a niche-transcending adaptation between ecotypes can result in homogeneity across ecotypes in the local chromosome neighborhood of the transferred segment [27]. (C) The transfer of a niche-neutral allele between ecotypes has no effect on the adaptive divergence between ecotypes. The frequencies of recombinant niche-neutral alleles follow the vagaries of genetic drift.

lowest carbon availability. Thus, the ability to function gluttonously as a copiotroph, in habitats where carbon is highly available, or to function frugally as an oligotroph in carbon-poor environments, has not evolved frequently. Similar studies have suggested that evolution of pH [24] and drought [25] tolerances are difficult transitions.

The Origins of Prokaryotic Species

While there may be scores of bauplans and extremely infrequent adaptations whose origins we would like to understand, there are likely many millions or even billions of individual species [7] whose origins beg our attention. After proposing an appropriate concept of species for prokaryotes, we will address how the population dynamics of prokaryotes affects their rates of speciation.

Concepts of Prokaryotic Species

The various modern concepts of species all attribute certain dynamic properties to species: that each species should be a cohesive group, whose diversity is limited by an evolutionary force; that different species are irreversibly separate; that species are ecologically distinct; and that species are each founded only once [26]. Efforts to define prokaryotic species according to these properties have differed most profoundly in the forces of cohesion deemed to be most important for prokaryotic species.

In the ecotype concept of species, a prokaryotic species (or ecotype) is a clade whose members are ecologically similar to one another, so that genetic diversity within the ecotype is limited by a cohesive force, either periodic selection or genetic drift, or both [27]. Periodic selection is the purging of diversity occurring when recombination is rare, such that natural selection favoring an adaptive mutation expunges diversity, genome-wide, within an ecotype. Alternatively, in bacterial populations of modest size (as with some pathogens, for example), diversity among members of an ecotype may be purged primarily by genetic drift. Owing to periodic selection or drift, the diversity within an ecologically homogeneous ecotype is only ephemeral. Divergence can become permanent when a mutation (or recombination event) places an organism into a new ecological niche and founds a new ecotype. Because the new ecotype is ecologically distinct from the parental ecotype, neither periodic selection nor drift events within the parental ecotype can extinguish the diversity within the new ecotype [27–29].

We have previously argued that ecotypes so defined are irreversibly separate from one another, regardless of the degree of resistance to recombination between ecotypes — sexual isolation [27,28,30] (Figure 2). Even if recombination were to occur between prokaryotic populations at the same rate as within them, it would not suffice to homogenize the adaptive, niche-specifying divergence between populations, owing to the extreme rarity of recombination in prokaryotes. Provided that the niche-specifying genes of each ecotype incur a cost when recombined into other ecotypes, natural selection will keep the frequency of recombinant, niche-specifying genes from other populations at negligible levels [27,28] (Figure 2). Recombination appears not to be a force that can halt or hinder adaptive, niche-specifying divergence between newly divergent ecotypes.

Recently, some microbiologists have become interested in recombination as a force of cohesion among closely related, ecologically distinct prokaryotic populations, and their models of speciation have focused on barriers to recombination [31–34]. For example, Sheppard and colleagues [33] claimed that two species of *Campylobacter* are 'despeciating' because recombination has recently transferred some niche-neutral sequence variants (used in multilocus sequence analysis and in phylogeny construction) from one species to another. It is important to recognize, however, that recombination has very different consequences for niche-neutral and niche-specifying variants (Figure 2) [35].

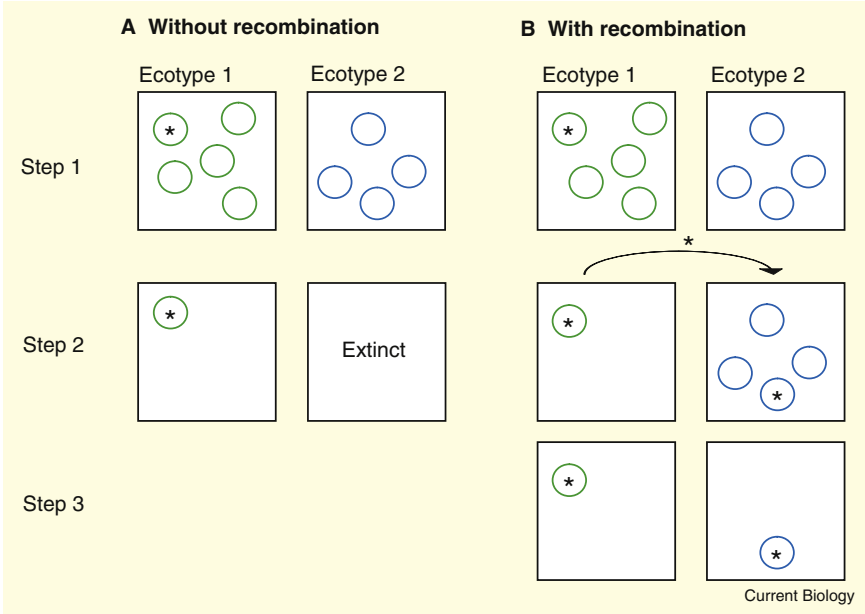

Figure 3. Recombination between nascent species can promote their coexistence.

Ecotypes 1 and 2 represent recently divergent ecotypes that differ only in the *proportions* of resources that they require. Because they have no unique resources, each ecotype may be vulnerable to exceptionally competitive mutants in the other. Circles represent different genotypes within each ecotype. Asterisks represent a mutation giving an exceptionally competitive advantage to the genotype that possesses it. (A) When no recombination occurs, the adaptive mutant in Ecotype 1 causes the extinction of Ecotype 2, along with all other genotypes in Ecotype 1. (B) With recombination, the adaptive allele may be transferred into Ecotype 2, causing a periodic selection event in that ecotype, and allowing for the continued coexistence of both ecotypes.

Because recombined niche-neutral variants can persist in a new ecotype, such recombination can be a nuisance *for us*, contributing to our inability to discover and classify ecotypes, but these recombination events are of no fitness consequence to the organism itself. Recombination may leave its mark on niche-neutral variation, but it is much too rare to cause homogenization of niche-specifying genes; even without sexual isolation, recombination cannot prevent the long-term coexistence of prokaryotic ecotypes as separate, ecologically distinct lineages [28]. Evolution of sexual isolation is therefore not a milestone of prokaryotic speciation, and it is not an appropriate criterion for demarcating prokaryotic species.

Ecotypes may then be defined as clades that are ecologically distinct, so that they escape each other's periodic selection and drift events, without concern for the degree of sexual isolation among them [27,28]. Ecotypes so defined bear all the dynamic properties attributed to species [26,27]: each ecotype is a cohesive group (whose diversity is limited by periodic selection and/or drift); different ecotypes are irreversibly separate (because they are out of range of one another's periodic selection and drift events, and because recombination is too rare to prevent their adaptive divergence); they are ecologically distinct (which allows them to coexist in the future); and ecotypes are founded only once.

High Net Rates of Speciation

The origins of ecotypes (or speciation) in prokaryotes may be accelerated by several features of prokaryotic population dynamics. First, because sexual isolation is not a prerequisite step in the origin of the species-like ecotypes of prokaryotes, we might expect that geographic isolation is not necessary in prokaryotic speciation, and we should expect to see examples of sympatric speciation in prokaryotes. One likely example is seen among most-closely-related ecotypes within *Bacillus simplex*, which are adapted to different solar exposures within canyons of northern Israel [36–39]; another example features most-closely-related ecotypes within *Vibrio splendidus*, which are adapted to particles of different sizes and to different times of year, within a Massachusetts estuary [40].

The enormous population sizes of many prokaryotes should make imminently accessible any adaptive mutation requiring only a single nucleotide substitution. Also, the most rare and unlikely of adaptive recombination events enter the realm of possibility.

Rare recombination between newly divergent ecotypes may reduce their rate of extinction [30] (Figure 3). At an early stage of speciation, a nascent ecotype may have diverged only in the *proportions* of different resources used, and utilize no novel resources of its own. In this case, newly divergent ecotypes may be vulnerable to extinction by periodic selection caused by an adaptive mutation within the parental ecotype. However, a single horizontal gene transfer event may transfer the adaptive mutation across ecotypes, and could thereby prevent extinction of one ecotype by another (while recombination is too rare to hinder niche-specifying divergence) [30]. It is ironic that recombination between nascent species, which hinders speciation in animals and plants, can facilitate the coexistence of young prokaryotic species [30].

Finally, the peculiar nature of genetic exchange in prokaryotes should promote speciation by allowing acquisition of heterologous genes from extremely distant relatives, as we discuss below.

Horizontal Gene Transfer as a Source of Ecological Innovation

Genome content comparisons shout to us a salient feature of adaptation — that prokaryotes can adapt to new environments and new ways of making a living by acquiring genes and sets of genes from other organisms [41]. Whereas each animal or plant lineage must invent each adaptation on its own, a prokaryotic lineage may acquire a pre-existing, niche-transcending adaptation — defined as adaptive in the genetic backgrounds of different ecotypes — from another lineage (Figure 2), and use it in the context of its own genetic background to create a novel niche.

Why is horizontal gene transfer so much more important in prokaryotes than in eukaryotes? One reason is that sexual isolation is much more extreme in animals and plants than

in prokaryotes [42], allowing adaptations in prokaryotes to arrive from more phylogenetically distant sources. Also, the rarity of genetic exchange in prokaryotes allows occasional, successful transfer of niche-transcending adaptations, without disturbing the integrity of niche-specifying adaptations (Figure 2) [28]. Lastly, recombined segments are short, allowing the transfer of a small, linked set of genes conferring a niche-transcending adaptation, without the co-transfer of niche-specifying genes [43].

Sudden acquisition by horizontal gene transfer of an entirely new physiological capability may present a challenge to the recipient organism. Animal and plant biologists have long understood the disruptive effects of sudden major genetic changes on the complex physiology and development of animals [44]. It might be argued that acquiring a new metabolic capacity might be less disruptive to the simple developmental and physiological system of a prokaryote [43].

Even if an horizontal gene transfer event (or a mutation of major effect) were grossly disruptive to a recipient prokaryote, the recombinant could still be successful [45]. The reason is that a new mutation or recombination event of major effect may bring a cell (and its clonal descendants) into a new ecological niche, where they may use a new set of resources not consumed by the parental ecotype. The criterion for success in this case is then not whether the cell can outcompete members of its previous ecotype; instead, the organism needs only to maintain a positive growth rate in its new niche.

Another factor facilitating adaptation by horizontal gene transfer is that the genes constituting a module of adaptation are frequently linked on the chromosome as an 'operon', where a set of functionally related genes are coded sequentially on the chromosome and are regulated together, as described by the 'selfish operon' model [16].

Changes in Existing Genes

Even if much of adaptation and invasion of new niches in prokaryotes is brought about by horizontal gene transfer, there is much opportunity for adaptation by changes in genes already existing in the genome. One reason is that transfer and incorporation of a new operon and biochemical pathway might create an imbalance in metabolism, creating new natural selection on pre-existing genes to restore balance; also, there may be natural selection on the horizontally acquired genes themselves to better accommodate the physiology of their new home [46]. Indeed, horizontally transferred genes have been shown to incur high rates of amino acid substitution after being transferred [47].

The invention of new niches can arise solely by changes in existing genes, as demonstrated by experiments in laboratory evolution [48]. In nature, mutations in existing genes appear to be at least partly responsible for evolution of new ecotypes, as seen in evolution of urovirulence in *Escherichia coli* [49], lung pathogenicity in *Pseudomonas aeruginosa* [50], host specificity in *Borrelia burgdorferi* [51], and defense against different amoebic predators in *Salmonella enterica* [52].

Additional evidence of mutation-based ecotype differences may eventually emerge from prokaryotic systematics. However, this will require systematics to change its focus from the presence-versus-absence of metabolic differences, which are most likely caused by acquisition and/or loss of genes, to *quantitative* differences in metabolic capabilities, which may yield evidence of evolution by mutational tweaking of existing genes [53].

Some adaptations are expected to be extremely easy to evolve by mutations in existing genes, particularly when a single nucleotide substitution can yield a significant increase in fitness, all the more likely when there are multiple mutations within a gene, or perhaps across different genes, that can bring about the adaptation [54,55]. Other changes are much more difficult to reach by mutation, for example when two separate mutations are necessary to reach an adaptation [56].

One limitation of evolution by mutation may be its incremental nature. It may not be possible to reach a particular adaptation because the intermediate evolutionary steps reached through a series of small mutations may be harmful. Nevertheless, botanists have discovered that a single mutation in a global regulatory protein, which changes the expression of many proteins simultaneously, may yield a startlingly novel morphology [57]. This evolutionary motif has occurred in natural populations of bacteria: many independent lineages of *Pseudomonas aeruginosa* have adapted to human lung pathogenicity by changes in the global regulator *lasR* [50].

Rates of Anagenesis and Cladogenesis in Prokaryotes

Evolutionary biologists have characterized the evolution of diversity into two categories: anagenesis, which is the accumulation of changes over time along a single lineage, and cladogenesis, which is the irreversible splitting of lineages [58]. Owing to the various obstacles to speciation in animals and plants, particularly the need for sexual isolation, the rate of anagenesis in animals and plants is much greater than that of cladogenesis.

Because speciation in prokaryotes does not require evolution of sexual isolation, one might expect that prokaryotic anagenesis and cladogenesis might be comparable in their rates, and we recently tested this hypothesis (A.K., J. Wertheim, L. Barone, N. Gentile, D. Krizanc, and F.C., unpublished data). We found that, in laboratory populations of *Bacillus subtilis*, new ecotypes consistently evolved before the original ecotype became fixed for even a single adaptive mutation, suggesting that speciation may be the predominant event of adaptive evolution in prokaryotes.

Global change toward warmer climates presents an interesting natural experiment regarding the relative rates of anagenesis and cladogenesis. We consider the various ecotypes within *Bacillus simplex* that are already specialized, through evolution of heat tolerance differences, to the north-facing or south-facing slopes of semi-arid canyons in northern Israel [36,39]. As the climate of northern Israel warms, each existing ecotype may possibly adapt through an anagenetic solution, simply increasing its own heat tolerance without creating any new ecotypes. Or adaptation may involve past cladogenesis, whereby already-existing south-facing ecotypes invade the north-facing slopes and displace the present north-facing ecotypes [59]. Lastly, adaptation to future global warming may involve invention of new ecotypes, which coexist with the parental ecotypes that spawn them.

Differences in Speciation Rates among Prokaryotic Groups

Consider next why some prokaryotic groups might be expected to undergo speciation at a faster rate than others. One possibility is that an ecological innovation may place a lineage into a new way of making a living, placing it within

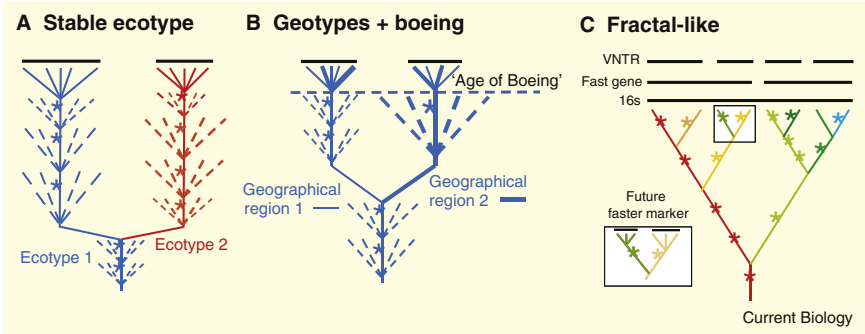

Figure 4. Models of prokaryotic diversification.

Ecotype formation events are represented by a change in color. Periodic selection events are represented by asterisks. Black lines above the phylogenies represent the sequence clusters likely to be demarcated as ecotypes by Ecotype Simulation. (A) The Stable Ecotype model, in which ecotype formation is relatively rare, and periodic selection events frequently purge all diversity from within each ecotype. This model predicts a one-to-one correspondence between sequence clusters and ecotypes. (B) The Geotypes + boeing model, in which prokaryotic populations separated by geography can belong to different sequence clusters even if they are of the same ecotype. This would predict a many-to-one relationship between sequence clusters and ecotypes, because the increased migration of prokaryotes brought about by recent advances in human travel could allow different sequence clusters of the same ecotype to exist in sympatry, at least temporarily. (C) The Fractal-like model, in which ecotypes are formed extremely rapidly, so that even the most rapidly evolving niche-neutral molecular markers fail to distinguish the most newly divergent ecotypes. Each molecular marker reveals ecologically distinct clusters that appear to be ecotypes, but more rapidly evolving markers frequently reveal even more newly divergent clusters that are ecologically distinct.

easy evolutionary access to many novel niches. For example, in the vertebrates, the invention of the desiccation-resistant amniote egg fostered adaptive radiation into many terrestrial habitats. Likewise, the evolution of the resilient membrane of the firmicutes may have promoted the adaptive radiation in this group into drought-prone habitats.

The evolution of developmental independence of different body parts is believed to have fostered rapid diversification in animals, by allowing separate optimization of different structures [60,61]. A prokaryotic analog may be independence of regulation of gene expression. For example, *Pseudomonas* has an extraordinarily high number of regulatory units, which may be responsible for the large number of disparate ecological niches held by closely related *Pseudomonas* species [62]. Additionally, successful integration of a newly acquired locus by horizontal gene transfer may require modulation of its expression [63], and luxuriant regulatory capacity may facilitate this modulation.

A higher rate of speciation may be expected when lineages are more finely specialized, as seen in moths and butterflies [64]. Likewise, we may predict higher rates of speciation in more specialized heterotrophic bacteria, for example in those organisms whose genomes bear fewer genes involved in uptake and processing of environmental carbon sources.

In some cases, ecological opportunity for rapid speciation may be afforded by the happenstance of living in a hard-to-reach region where many competing species are not present, as seen in the zoological and botanical adaptive radiations on oceanic islands. While some bacteria are seasoned worldwide travelers, for example spore-forming bacteria such as *Bacillus*, there appear to be constraints on the migration of many prokaryotic groups [65], and so oceanic islands may provide a crucible for bacterial adaptive radiation in these groups.

The Ecological Dimensions of Prokaryotic Speciation

Identifying Newly Divergent Ecotypes

Recognizing newly divergent ecotypes will help to determine the ecological dimensions by which prokaryotes diverge upon speciation and by which they may coexist. One might imagine that comparing closely related, named species would provide insight into these first ecological differences in cladogenesis. However, the species of prokaryotic systematics are extremely broadly defined, as measured by any criterion of difference, including divergence of orthologous sequences [66], genome content and physiology [67], and most importantly, ecology [68]. The named species of prokaryotic systematics are therefore demarcated more like a genus than a species, at least as defined in animals or plants. If we are to identify the first ecological changes in newly divergent lineages, we have to identify a taxon more narrowly defined than the species of prokaryotic systematics.

There has been growing interest in discovering ecotypes within the recognized species, and they have frequently been observed [36,38,69–74]. Moreover, interest is growing for a more ecologically based systematics, where ecotypes are considered the fundamental units of diversity [27,34,36,75–79]. A special challenge to identifying these ecotypes is that an ecologist will not generally be able to anticipate the phenotypic characters that determine the ecological distinctness of as yet undiscovered ecotypes. This is because invention of new ecotypes may frequently involve the chance acquisition of genes from other taxa, sometimes from extremely divergent taxa, and one cannot predict the genes and functions that will be involved, even in well studied taxa [27].

Longstanding ecotypes may in principle be discovered and classified through DNA sequence analysis [27,36]. This will be the case under the 'Stable Ecotype' model of prokaryotic evolution, in which ecotypes are formed rarely enough so that each ecotype has time enough to accumulate its own unique set of sequence mutations, while diversity within ecotypes is recurrently purged by periodic selection and/or drift, yielding a correspondence between ecotypes and sequence clusters for any gene shared among ecotypes (Figure 4A).

This approach has been used for decades as a means for discovering ecological diversity in prokaryotes, and much ecological diversity within named species has been discovered in this way [80]. However, it is difficult to identify ecotypes through sequence clusters, as any phylogenetic tree contains a hierarchy of subclusters within clusters, and it is generally not clear which level of sequence cluster corresponds to ecologically distinct populations. This challenge has motivated the recent development of two computer algorithms that simulate the evolutionary dynamics of

Table 1. Algorithms that hypothesize ecotype demarcations independently from any *a priori*, universal cutoff. Each demarcates on the basis of the best fit to a dynamic model of the evolutionary history of a clade.

Algorithm	Input	Output	Advantages	Disadvantages
Ecotype simulation [36]	Recombination-free concatenation of multilocus sequence data	Rates of drift, periodic selection, and ecotype formation	Ecotype demarcation is not limited to ecotypes that have diverged on a specified set of ecological dimensions	Hypothesized ecotypes must be confirmed independently as ecologically distinct
AdaptML [40]	Phylogeny (based on any data); habitat source of each organism	Rate of ecotype formation; quantification of ecotypes' habitat preferences	Hypothesized ecotypes are confirmed by the algorithm as ecologically distinct	Detects only those ecotypes that are ecologically distinct on the ecological dimensions specified[1]

[1]AdaptML can detect only those ecotypes that are distinct for the habitat variables specified in input. For example, this led to the conclusion that no adaptive radiation occurred within a clade containing *V. calviensis* and *Enterovibrio norvegicus* [40]. However, these species are each nearly as phylogenetically diverse as *V. splendidus*, where an adaptive radiation was shown to have occurred, and an Ecotype Simulation analysis of the *calviensis-norvegicus* clade containing these two species indicated seven ecotypes (unpublished data). This clade may have undergone adaptive radiations in environmental variables not targeted by the investigators.

ecotypes to hypothesize ecotype demarcations from sequence data: Ecotype Simulation [36] and AdaptML [40]. Each analyzes the evolutionary history of a particular clade to yield the appropriate criteria for demarcating ecotypes (Table 1).

Another significant challenge to sequence-based discovery of ecotypes is to take into account the circumstances where sequence clusters do *not* correspond to ecotypes. For example, there may be multiple, longstanding sequence clusters within a single ecotype when there has been a history of geographic isolation [27] (Figure 4B). Therefore, recognition of ecotypes requires confirmation that the putative ecotypes hypothesized by sequence analysis are ecologically distinct [36].

Ecological distinctness of putative ecotypes can be readily confirmed when different putative ecotypes are associated with different microhabitats or hosts. We therefore encourage microbial ecologists and systematists to sample across contrasting microhabitats. Also, inclusion of detailed environmental data with submission of sequences to GenBank will allow future ecologists and systematists to discover and confirm ecotypes from sequence data.

The ecotypes hypothesized by Ecotype Simulation, even those within the usual phylogenetic range of a named species — for example, those with less than 1% divergence in 16S rRNA [81] — have generally been confirmed as different in their ecology. Putative ecotypes identified in arid-soil *Bacillus* have differed in their adaptations to solar exposure [36], as well as in adaptations to gravel versus clay soil [82]. Very closely related putative ecotypes within hot spring *Synechococcus* differ in adaptation to temperature and depth in the photic zone [83], and ecotypes within *Legionella pneumophila* differ in their host range and gene expression patterns [74]. Ecotype Simulation promises to be an effective means of discovering very closely related ecotypes, as does AdaptML (Table 1).

We next consider how ecotypes might be discovered and confirmed when the ecological dimensions of their adaptive divergence are beyond the imagination of ecologists. Here microbial ecology may derive inspiration from a biotic approach for characterizing environments, developed long ago by zoologists and botanists [84]. Each site where a focus clade of prokaryotes is collected could be biotically characterized by the whole of the prokaryotic community that lives there, using a survey of 16S rRNA from the community's environmental DNA [85]. The ecological distinctness of putative ecotypes within a focus clade could be confirmed by association with different, biotically-defined types of communities. Thus, microbial ecologists may move all the way from hypothesizing ecotypes to confirming their ecological distinctness, without requiring any *a priori* knowledge of the ecological dimensions of prokaryotic speciation. Moreover, the discovery that a given ecotype is associated with a particular biotic community may eventually yield information about the conditions and resources required by the ecotype. For example, knowing that an ecotype is associated with the Bacteriodetes would suggest a copiotrophic lifestyle. As more detailed information about taxon lifestyles become known, associations with biotic communities will become more informative.

Consider next how to discover the physiological and genomic adaptations that are responsible for ecological differences. Differences in gene content, as revealed in genome comparisons, have revealed ecological differences among closely related pathogens [86] and free-living phototrophs [87], as well as in other groups. Indeed, the newly launched Genomic Encyclopedia of Bacteria and Archaea (GEBA), aiming to sequence the type strain of each named species of prokaryote, is expected to provide new ways of discovering ecological differences from genome content differences [88]. In addition, differences in genome-wide gene expression [68] and comprehensive metabolic analysis [89] may reveal the mechanisms of ecological distinctness.

Any ecology-blind, sequence-based discovery system is limited by the resolution of the molecular markers used, and will not detect ecotypes too newly invented to have accrued neutral molecular divergence. We suspect that a much faster molecular marker, such as the variable number of tandem repeats (VNTR) loci [90] may discern newly invented ecotypes that are subsumed within a single taxon distinguishable by nucleotide substitutions. We hypothesize that the ecological diversity in some clades of prokaryotes may be almost fractal-like, with younger and younger ecologically distinct clades waiting to be discovered by ever faster molecular markers (Figure 4C).

Considering that the vast majority of prokaryotes are currently uncultivable, a more general approach toward identifying ecological differences among close relatives will have to be independent of isolation and cultivation of organisms from nature. To this end, metagenomics — high-throughput sequencing and assembly of random segments of uncultivated DNA from an entire community — may provide a systematic and cultivation-free approach to discovering adaptations that distinguish closest relatives [68]. One rationale

is to find niche-specifying genes — for example, coding for ability to use an environmental substrate — whose presence is variable among close relatives. Bhaya *et al.* [87] performed such an analysis on environmental sequences closely related to one isolate of hot spring *Synechococcus*, and found that close relatives of the isolate frequently differed in presence-versus-absence of a ferrous-uptake pathway. This indicates that this is a recently evolved (or at least recently acquired) adaptation in this clade.

The Everyday Transitions of Prokaryotic Ecology

What are the ecological transitions seen between the most closely related ecotypes? In the case of pathogens, a change in the host species is frequently observed. Systematics has recognized pathovars adapted to alternative hosts within several species, for example within *Xanthomonas campestris* [91], but in other cases host-specific clades within named bacterial species are not recognized by systematics [51,73,92]. Closely related mutualists may also differ in their host ranges, as seen in *Rhizobium* [93].

Ecotypes of closely related pathogens may also differ in the tissues they infect, which can change the mode of transmission, as seen in *Treponema* [94], *Streptococcus* [70], and *Wolbachia* [92]. Some pathogens have an extended environmental phase as they pass between hosts, yielding an opportunity for ecotypes to diverge in their adaptations to environmental conditions [95].

Perhaps the easiest ecological transitions are those that occur repeatedly within a species. For example, in various bacterial species, a particular virulence adaptation may appear within an individual human host, leading to an ephemeral, virulent lineage that dies within the individual host [96]; such adaptations may occur repeatedly within a species, as seen in *Haemophilus influenzae* [97] and in *Pseudomonas aeruginosa* [50].

In the free-living heterotrophs, closely related ecotypes may differ along a number of environmental dimensions, including the carbon sources that they can use [71]. Closely related heterotrophic ecotypes may differ in adaptations to physical conditions, such as temperature [36,38,39] and pH [98], and the size of soil [82] and marine [40] particles, as well as season of activity [40].

Beyond the ecotypes that we can pin to known ecological dimensions, there exist many as yet uncharacterized specializations among closely related, heterotrophic ecotypes. For example, several microgeographic studies have found an association between closeness of relationship (within a named species) and closeness of geography within a small region [99–101]. These associations were likely caused by specializations to unknown environmental variables, since migration effects could be ruled out. In other cases, multiple closely related ecotypes from different clades appear specialized to the same conditions. For example, the various *B. simplex* ecotypes specialized to the north-facing slopes of arid canyons in northern Israel likely owe their coexistence to divergence in unknown environmental conditions or resources. These examples suggest that in heterotrophs there may be a world of unknown dimensions to ecological divergence.

Ecological divergence in the prokaryotic phototrophs presents the same fundamental challenge that has vexed plant ecologists — explaining how a diversity of phototrophs so superficially similar in their resource demands can manage to partition resources and thereby coexist. Extremely close relatives [83] and more distant relatives [69] within the cyanobacteria have specialized to alternative depths (and photic levels). Closely related cyanobacteria may also differ in the ions used to obtain a given nutrient element, such as nitrite versus ammonium between closely related marine *Prochlorococcus* [102], and phosphate versus phosphonate in hot spring *Synechococcus* [87]. Closely related phototrophs may be specialized to different temperatures [83], and perhaps related to temperature adaptation is specialization to different seasons as seen in marine *Prochlorococcus* [103].

Beyond adapting to the resources and physical conditions of a particular habitat, an organism must adapt to the organisms that are present, as is well known in plants and animals [104]. In the case of bacteria, predators may determine the geographical distribution of prey ecotypes. For example, the various O antigens of *Salmonella* each defend against a different species of amoebic predator [52]. Bacteriophage may possibly foster coexistence of closely related bacteria by most efficiently transmitting within host populations at highest densities [105].

There is also much potential for interference competition among closely related ecotypes through release of antagonistic bacteriocin compounds [106]. More generally, Davies and colleagues [107] have shown that there is a great potential for modulation of bacterial physiology by small molecules secreted by other bacteria. Given the wide diversity of organisms (and their secretions) that could affect a given organism, it is not clear how much of an ecotype's adaptation to a given habitat involves accommodation to the habitat's resources and physical conditions, and how much involves accommodation to predators and secreted substances.

Discovery of ecotypes provides our best window on the early ecological divergences that can set in motion the long-term, irreversible divergence between lineages. Understanding the ecological bases of ecotype differentiation also helps us understand how partitioning of resources, as well as differences in adaptation to physical conditions and other organisms, can explain the coexistence of the many species within a prokaryotic clade.

Discovery of extensive ecotype diversity within recognized species should increase our estimates of prokaryotic diversity, including our estimates of rare and endemic taxa. Even for world-travelling groups such as *Bacillus*, the discovery of infraspecific ecotypes adapted to endemic animals and plants should lead to higher estimates of bacterial endemism.

Recommendations

Leveraging What We Already Know and What We Can Find Out

Prokaryotic systematics is squeezed between the enormity of its task to characterize a billion species, as well as 100 disparate divisions, and the difficulty of discovering and characterizing even a single species. We are profoundly challenged by the lack of *a priori* information on the physiological features likely responsible for adaptive divergence between closely related species, and by often not even knowing the ecological dimensions by which nascent species are most likely to diverge. To discover the full extent of prokaryotic diversity, even within a single clade in a single community, we have to develop ways to leverage what we already know and what we can most easily find out.

Systematics can overcome its lack of physiological and ecological information about species divergence through ecology-blind algorithms for discovering and demarcating ecotypes (for example, Ecotype Simulation). Putative ecotypes so demarcated can be confirmed by testing for associations with habitat types that are hypothesized by the investigators. Requiring less dependence on ecological insight, putative ecotypes may be alternatively confirmed as ecologically distinct through associations with different, biotically defined communities.

To better leverage our insight and our data, ecologists will need to more fully measure and submit habitat data when they submit sequence data to on-line sequence repositories. GenBank allows for publication of such habitat data, although there is not yet a standardized set of environmental parameters. For soil habitats, this effort may begin with parameters such as salinity, water potential, and organic content, but the parameter set should grow as investigators discover new environmental parameters that represent important ecological dimensions of speciation (for example, solar exposure).

Such a habitat database would allow future investigators to identify and confirm ecotypes, perhaps through Ecotype Simulation [36] and AdaptML [40]. It would also allow identification of the easiest and most difficult of habitat transitions [22], even when the differences between habitats are subtle. When applied to biotic classification of habitats, a habitat database would also allow determination of the environmental preferences of each of many thousands of taxa, allowing ecologists to move all the way from identification of ecotypes within a focus clade, to confirmation of the putative ecotypes' distinctness, and finally to characterize the ecotypes' habitat differences, all without any *a priori* knowledge of the ecological dimensions of speciation of the focus clade.

Metagenomics can help systematists around the handicap of not knowing the physiological differences to expect among ecotypes, by providing a list of likely niche-specifying genes that are present and absent in different individuals within a clade [87]. After isolates within a particular clade are demarcated into putative ecotypes — for example, based on Ecotype Simulation analysis of some shared genes — the ecotypes could then be tested for presence of the niche-specifying genes suggested by metagenomics. This targeted analysis of genome content would circumvent the need for full-genome sequencing of every putative ecotype.

Genome sequencing of a single cell, without cultivation [108], promises to reveal the physiological and ecological differences between organisms of every taxonomic level, from nascent ecotypes to anciently divergent divisions. This represents a profound leverage of inference from well studied organisms to the utterly unfamiliar. However, we do not know the functions of many genes in any genome, nor do we know the genes responsible for many ecological adaptations, for example salt tolerance. Ecological inference from genomics will be an extremely important goal, and fortunately some general solutions are on the horizon [109].

Lastly, a profound handicap of prokaryotic systematics is that it does not attempt to incorporate ecology into species demarcations, and in many cases multiple ecotypes are subsumed within a single species taxon. To derive the full benefits of identifying ecotypes, systematics should recognize ecotypes as valid taxa, as we have previously suggested [27,36].

Acknowledgments

We are grateful to Dana Hunt and Martin Polz for sharing their marine *Vibrio* sequence data. This work was supported by NSF grant EF-0328698 and by research funds from Wesleyan University.

References

1. Conway Morris, S. (1995). A new phylum from the lobster's lips. Nature *378*, 661–662.
2. Giovannoni, S.J., and Stingl, U. (2005). Molecular diversity and ecology of microbial plankton. Nature *437*, 343–348.
3. Woese, C.R. (1987). Bacterial evolution. Microbiol. Rev. *51*, 221–271.
4. May, R.M. (1988). How many species are there on earth? Science *241*, 1441–1449.
5. Euzéby, J.P. (2008). List of prokaryotic names with standing in nomenclature. www.bacterio.cict.fr/number.html#total
6. Dykhuizen, D.E. (1998). Santa Rosalia revisited: why are there so many species of bacteria? Antonie Van Leeuwenhoek *73*, 25–33.
7. Gans, J., Wolinsky, M., and Dunbar, J. (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science *309*, 1387–1390.
8. Madigan, M.T., Martinko, J.M., and Parker, J. (2000). Brock Biology of Microorganisms, 9th Edition, (Upper Saddle River, NJ: Prentice-Hall).
9. Esteve-Nuñez, A., Lucchesi, G., Philipp, B., Schink, B., and Ramos, J.L. (2000). Respiration of 2,4,6-trinitrotoluene by *Pseudomonas* sp. strain JLR11. J. Bacteriol. *182*, 1352–1355.
10. W. Arthur, ed. (2006). Creatures of Accident: The Rise of the Animal Kingdom (New York: Hill and Wang), p. 54.
11. Knoll, A.H. (2003). Life on a Young Planet: The First Three Billion Years of Evolution on Earth (Princeton: Princeton University Press).
12. Summons, R.E., Bradley, A.S., Jahnke, L.L., and Waldbauer, J.R. (2006). Steroids, triterpenoids and molecular oxygen. Philos. Trans. R. Soc. Lond. B Biol. Sci. *361*, 951–968.
13. Koonin, E.V. (2007). The Biological Big Bang model for the major transitions in evolution. Biol. Direct. *2*, 21.
14. Szathmáry, E., and Maynard Smith, J. (1995). The Major Transitions in Evolution (Oxford: Oxford University Press).
15. Norberg, U.M. (1985). Evolution of vertebrate flight: an aerodynamic model for the transition from gliding to active flight. Am. Nat. *126*, 303–327.
16. Lawrence, J.G. (1997). Selfish operons and speciation by gene transfer. Trends. Microbiol. *5*, 355–359.
17. Liu, R., and Ochman, H. (2007). Stepwise formation of the bacterial flagellar system. Proc. Natl. Acad. Sci. USA *104*, 7116–7121.
18. Vitt, L.J., and Pianka, E.R. (2005). Deep history impacts present-day ecology and biodiversity. Proc. Natl. Acad. Sci. USA *102*, 7877–7881.
19. Daniel, R.M., and Cowan, D.A. (2000). Biomolecular stability and life at high temperatures. Cell Mol. Life. Sci. *57*, 250–264.
20. Schimel, J., Balser, T.C., and Wallenstein, M. (2007). Microbial stress-response physiology and its implications for ecosystem function. Ecology *88*, 1386–1394.
21. Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. Appl. Environ. Microbiol. *71*, 8228–8235.
22. Lozupone, C.A., and Knight, R. (2007). Global patterns in bacterial diversity. Proc. Natl. Acad. Sci. USA *104*, 11436–11440.
23. Fierer, N., Bradford, M.A., and Jackson, R.B. (2007). Toward an ecological classification of soil bacteria. Ecology *88*, 1354–1364.
24. Fierer, N., and Jackson, R.B. (2006). The diversity and biogeography of soil bacterial communities. Proc. Natl. Acad. Sci. USA *103*, 626–631.
25. Nazih, N., Finlay-Moore, O., Hartel, P.G., and Fuhrmann, J.J. (2001). Whole soil fatty acid methyl ester (FAME) profiles of early soybean rhizosphere as affected by temperature and matric water potential. Soil Biol. Biochem. *33*, 693–696.
26. de Queiroz, K. (2005). Ernst Mayr and the modern concept of species. Proc. Natl. Acad. Sci. USA *102 (Suppl 1)*, 6600–6607.
27. Cohan, F.M., and Perry, E.B. (2007). A systematics for discovering the fundamental units of bacterial diversity. Curr. Biol. *17*, R373–R386.
28. Cohan, F.M. (1994). The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. Am. Nat. *143*, 965–986.
29. Templeton, A. (1989). The meaning of species and speciation: a genetic perspective. In Speciation and its Consequences, D. Otte and J. Endler, eds. (Sunderland MA: Sinauer Assoc), pp. 3–27.
30. Cohan, F.M. (2001). Bacterial species and speciation. Syst. Biol. *50*, 513–524.
31. Fraser, C., Hanage, W.P., and Spratt, B.G. (2007). Recombination and the nature of bacterial speciation. Science *315*, 476–480.
32. Papke, R.T., Zhaxybayeva, O., Feil, E.J., Sommerfeld, K., Muise, D., and Doolittle, W.F. (2007). Searching for species in haloarchaea. Proc. Natl. Acad. Sci. USA *104*, 14092–14097.

33. Sheppard, S.K., McCarthy, N.D., Falush, D., and Maiden, M.C. (2008). Convergence of *Campylobacter* species: implications for bacterial evolution. Science *320*, 237–239.

34. Achtman, M., and Wagner, M. (2008). Microbial diversity and the genetic nature of microbial species. Nat. Rev. Microbiol. *6*, 431–440.

35. Bohannon, J., (2008). The case of the half-blood bug. In ScienceNOW Daily News.

36. Koeppel, A., Perry, E.B., Sikorski, J., Krizanc, D., Warner, W.A., Ward, D.M., Rooney, A.P., Brambilla, E., Connor, N., Ratcliff, R.M., *et al*. (2008). Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. Proc. Natl. Acad. Sci. *105*, 2504–2509.

37. Sikorski, J., Brambilla, E., Kroppenstedt, R.M., and Tindall, B.J. (2008). The temperature-adaptive fatty acid patterns in *Bacillus simplex* from "Evolution Canyons" I and II, Israel. Microbiology *154*, 2416–2426.

38. Sikorski, J., and Nevo, E. (2005). Adaptation and incipient sympatric speciation of *Bacillus simplex* under microclimatic contrast at "Evolution Canyons" I and II, Israel. Proc. Natl. Acad. Sci. USA *102*, 15924–15929.

39. Sikorski, J., and Nevo, E. (2007). Patterns of thermal adaptation of *Bacillus simplex* to the microclimatically contrasting slopes of 'Evolution Canyons' I and II, Israel. Environ. Microbiol. *9*, 716–726.

40. Hunt, D.E., David, L.A., Gevers, D., Preheim, S.P., Alm, E.J., and Polz, M.F. (2008). Resource partitioning and sympatric differentiation among closely related bacterioplankton. Science *320*, 1081–1085.

41. Ochman, H., and Davalos, L.M. (2006). The nature and dynamics of bacterial genomes. Science *311*, 1730–1733.

42. Cohan, F.M. (2002). Sexual isolation and speciation in bacteria. Genetica *116*, 359–370.

43. Zawadzki, P., and Cohan, F.M. (1995). The size and continuity of DNA segments integrated in Bacillus transformation. Genetics *141*, 1231–1243.

44. Fisher, R.A. (1958). The Genetical Theory of Natural Selection (New York: Dover).

45. Arthur, W. (1984). Mechanisms of Morphological Evolution: A Combined Genetic, Developmental and Ecological Approach (New York: Wiley & Sons), pp. 182–186.

46. Cohan, F.M., King, E.C., and Zawadzki, P. (1994). Amelioration of the deleterious pleiotropic effects of an adaptive mutation in *Bacillus subtilis*. Evolution *48*, 81–95.

47. Hao, W., and Golding, G.B. (2006). The fate of laterally transferred genes: life in the fast lane to adaptation or death. Genome Res. *16*, 636–643.

48. Bantinaki, E., Kassen, R., Knight, C.G., Robinson, Z., Spiers, A.J., and Rainey, P.B. (2007). Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader diversity. Genetics *176*, 441–453.

49. Sokurenko, E.V., Chesnokova, V., Dykhuizen, D.E., Ofek, I., Wu, X.R., Krogfelt, K.A., Struve, C., Schembri, M.A., and Hasty, D.L. (1998). Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesin. Proc. Natl. Acad. Sci. USA *95*, 8922–8926.

50. D'Argenio, D.A., Wu, M., Hoffman, L.R., Kulasekara, H.D., Deziel, E., Smith, E.E., Nguyen, H., Ernst, R.K., Larson Freeman, T.J., Spencer, D.H., *et al*. (2007). Growth phenotypes of *Pseudomonas aeruginosa lasR* mutants adapted to the airways of cystic fibrosis patients. Mol. Microbiol. *64*, 512–533.

51. Brisson, D., and Dykhuizen, D.E. (2004). *ospC* diversity in *Borrelia burgdorferi*: different hosts are different niches. Genetics *168*, 713–722.

52. Wildschutte, H., and Lawrence, J.G. (2007). Differential *Salmonella* survival against communities of intestinal amoebae. Microbiology *153*, 1781–1789.

53. Feldgarden, M., Byrd, N., and Cohan, F.M. (2003). Gradual evolution in bacteria: evidence from *Bacillus* systematics. Microbiology *149*, 3565–3573.

54. Jin, D.J., and Gross, C.A. (1989). Characterization of the pleiotropic phenotypes of rifampin-resistant *rpoB* mutants of *Escherichia coli*. J. Bacteriol. *171*, 5229–5231.

55. Albert, T.J., Dailidiene, D., Dailide, G., Norton, J.E., Kalia, A., Richmond, T.A., Molla, M., Singh, J., Green, R.D., and Berg, D.E. (2005). Mutation discovery in bacterial genomes: metronidazole resistance in *Helicobacter pylori*. Nat. Methods *2*, 951–953.

56. Blount, Z.D., Borland, C.Z., and Lenski, R.E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Proc. Natl. Acad. Sci. USA *105*, 7899–7906.

57. Sangster, T.A., Salathia, N., Undurraga, S., Milo, R., Schellenberg, K., Lindquist, S., and Queitsch, C. (2008). HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. Proc. Natl. Acad. Sci. USA *105*, 2963–2968.

58. Futuyma, D.J. (1987). On the role of species in anagenesis. Am. Nat. *130*, 465–473.

59. Cohan, F.M. (2008). Tracking bacterial responses to global warming with an ecotype-based systematics. Clin Microbiol. Infection, in press.

60. Vermeij, G.J. (1973). Adaptation, versatility, and evolution. Syst. Zool. *22*, 466–477.

61. Minelli, A. (2008). Perspectives on Animal Phylogeny and Evolution (Oxford: Oxford University Press).

62. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., *et al*. (2000). Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. Nature *406*, 959–964.

63. Doyle, M., Fookes, M., Ivens, A., Mangan, M.W., Wain, J., and Dorman, C.J. (2007). An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. Science *315*, 251–252.

64. Dyer, L.A., Singer, M.S., Lill, J.T., Stireman, J.O., Gentry, G.L., Marquis, R.J., Ricklefs, R.E., Greeney, H.F., Wagner, D.L., Morais, H.C., *et al*. (2007). Host specificity of Lepidoptera in tropical and temperate forests. Nature *448*, 696–699.

65. Papke, P.T., and Ward, D.M. (2004). The importance of physical isolation in microbial evolution. FEMS Microbiol. Ecol. *48*, 293–303.

66. Staley, J.T. (2006). The bacterial species dilemma and the genomic-phylogenetic species concept. Philos. Trans. R. Soc. Lond. B Biol. Sci. *361*, 1899–1909.

67. Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and Tiedje, J.M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. *57*, 81–91.

68. Ward, D.M., Cohan, F.M., Bhaya, D., Heidelberg, J.F., Kuhl, M., and Grossman, A. (2008). Genomics, environmental genomics and the issue of microbial species. Heredity *100*, 207–219.

69. Coleman, M.L., Sullivan, M.B., Martiny, A.C., Steglich, C., Barry, K., Delong, E.F., and Chisholm, S.W. (2006). Genomic islands and the ecology and evolution of *Prochlorococcus*. Science *311*, 1768–1770.

70. Enright, M.C., Spratt, B.G., Kalia, A., Cross, J.H., and Bessen, D.E. (2001). Multilocus sequence typing of *Streptococcus pyogenes* and the relationships between emm type and clone. Infect. Immun. *69*, 2416–2427.

71. Höfle, M.G., Ziemke, F., and Brettar, I. (2000). Niche differentiation of *Shewanella putrefaciens* populations from the Baltic as revealed by molecular and metabolic fingerprinting. In Microbial Biosystems: New Frontiers, Proceedings of the 8th International Symposium on Microbial Ecology, C.R. Bell, M. Brylinsky, and P. Johnson-Green, eds. (Halifax: Atlantic Canada Society for Microbial Ecology), pp. 135–142.

72. Paraskevopoulos, C., Bordenstein, S.R., Wernegreen, J.J., Werren, J.H., and Bourtzis, K. (2006). Toward a *Wolbachia* multilocus sequence typing system: discrimination of *Wolbachia* strains present in *Drosophila* species. Curr. Microbiol. *53*, 388–395.

73. Smith, N.H., Kremer, K., Inwald, J., Dale, J., Driscoll, J.R., Gordon, S.V., van Soolingen, D., Hewinson, R.G., and Smith, J.M. (2006). Ecotypes of the *Mycobacterium tuberculosis* complex. J. Theor. Biol. *239*, 220–225.

74. Cohan, F.M., Koeppel, A., and Krizanc, D. (2006). Sequence-based discovery of ecological diversity within *Legionella*. In Legionella: State of the Art 30 Years after Its Recognition, N.P. Cianciotto, Y. Abu Kwaik, P.H. Edelstein, B.S. Fields, D.F. Geary, T.G. Harrison, C. Joseph, R.M. Ratcliff, J.E. Stout, and M.S. Swanson, eds. (Washington, D.C.: ASM Press), pp. 367–376.

75. Konstantinidis, K.T., and Tiedje, J.M. (2007). Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. Curr. Opin. Microbiol. *10*, 504–509.

76. Palys, T., Nakamura, L.K., and Cohan, F.M. (1997). Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. Int. J. Syst. Bacteriol. *47*, 1145–1156.

77. Ward, D.M. (1998). A natural species concept for prokaryotes. Curr. Opin. Microbiol. *1*, 271–277.

78. Dellaglio, F., Felis, G.E., and Germond, J.E. (2004). Should names reflect the evolution of bacterial species? Int. J. Syst. Evol. Microbiol. *54*, 279–281.

79. Sikorski, J. (2008). Populations under microevolutionary scrutiny: what will we gain? Arch. Microbiol. *189*, 1–5.

80. Ahlgren, N.A., and Rocap, G. (2006). Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. Appl. Environ. Microbiol. *72*, 7193–7204.

81. Stackebrandt, E., and Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. Microbiol. Today *33*, 152–155.

82. Connor, N. (2008). Sequence-based discovery of ecological diversity in Death Valley *Bacillus* In Biology, M.A. (Middletown, Conn: Wesleyan University).

83. Ward, D.M., Bateson, M.M., Ferris, M.J., Kühl, M., Wieland, A., Koeppel, A., and Cohan, F.M. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. Phil. Trans. Roy. Soc. Ser. B. *361*, 1997–2008.

84. Gauch, H.G. (1982). Multivariate Analysis in Community Ecology (Cambridge: Cambridge University Press).

85. Lozupone, C.A., and Knight, R. (2008). Species divergence and the measurement of microbial diversity. FEMS. Microbiol. Rev. *32*, 557–578.

86. Guidot, A., Prior, P., Schoenfeld, J., Carrere, S., Genin, S., and Boucher, C. (2007). Genomic structure and phylogeny of the plant pathogen *Ralstonia solanacearum* inferred from gene distribution analysis. J. Bacteriol. *189*, 377–387.

87. Bhaya, D., Grossman, A.R., Steunou, A.S., Khuri, N., Cohan, F.M., Hamamura, N., Melendrez, M.C., Bateson, M.M., Ward, D.M., and Heidelberg,

J.F. (2007). Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. Isme. J. *1*, 703–713.

88. Liolios, K., Mavromatis, K., Tavernarakis, N., and Kyrpides, N.C. (2008). The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res. *36*, D475–D479.
89. Bochner, B.R. (2003). New technologies to assess genotype-phenotype relationships. Nat. Rev. Genet. *4*, 309–314.
90. Kenefic, L.J., Beaudry, J., Trim, C., Daly, R., Parmar, R., Zanecki, S., Huynh, L., Van Ert, M.N., Wagner, D.M., Graham, T., *et al*. (2008). High resolution genotyping of *Bacillus anthracis* outbreak strains using four highly mutable single nucleotide repeat markers. Lett. Appl. Microbiol. *46*, 600–603.
91. Van den Mooter, M., and Swings, J. (1990). Numerical analysis of 295 phenotypic features of 266 *Xanthomonas* strains and related strains and an improved taxonomy of the genus. Int. J. Syst. Bacteriol. *40*, 348–369.
92. Baldo, L., Dunning Hotopp, J.C., Jolley, K.A., Bordenstein, S.R., Biber, S.A., Choudhury, R.R., Hayashi, C., Maiden, M.C., Tettelin, H., and Werren, J.H. (2006). Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. Appl. Environ. Microbiol. *72*, 7098–7110.
93. Wernegreen, J.J., Harding, E.E., and Riley, M.A. (1997). *Rhizobium* gone native: unexpected plasmid stability of indigenous *Rhizobium leguminosarum*. Proc. Natl. Acad. Sci. USA *94*, 5483–5488.
94. Harper, K.N., Ocampo, P.S., Steiner, B.M., George, R.W., Silverman, M.S., Bolotin, S., Pillay, A., Saunders, N.J., and Armelagos, G.J. (2008). On the origin of the treponematoses: a phylogenetic approach. PLoS Neglected Tropical Diseases *2*, 1–13.
95. Carvalho, F.R., Vazoller, R.F., Foronda, A.S., and Pellizari, V.H. (2007). Phylogenetic study of *Legionella* species in pristine and polluted aquatic samples from a tropical Atlantic forest ecosystem. Curr. Microbiol. *55*, 288–293.
96. Levin, B.R., and Bull, J.J. (1994). Short-sighted evolution and the virulence of pathogenic microorganisms. Trends Microbiol. *2*, 76–81.
97. Margolis, E., and Levin, B.R. (2007). Within-host evolution for the invasiveness of commensal bacteria: an experimental study of bacteremias resulting from *Haemophilus influenzae* nasal carriage. J. Infect. Dis. *196*, 1068–1075.
98. Newton, R.J., Jones, S.E., Helmus, M.R., and McMahon, K.D. (2007). Phylogenetic ecology of the freshwater *Actinobacteria* acI lineage. Appl. Environ. Microbiol. *73*, 7169–7176.
99. Belotte, D., Curien, J.B., Maclean, R.C., and Bell, G. (2003). An experimental test of local adaptation in soil bacteria. Evolution *57*, 27–36.
100. Vos, M., and Velicer, G.J. (2008). Isolation by distance in the spore-forming soil bacterium *Myxococcus xanthus*. Curr. Biol. *18*, 386–391.
101. Bent, S.J., Gucker, C.L., Oda, Y., and Forney, L.J. (2003). Spatial distribution of *Rhodopseudomonas palustris* ecotypes on a local scale. Appl. Environ. Microbiol. *69*, 5192–5197.
102. Rocap, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N.A., Arellano, A., Coleman, M., Hauser, L., Hess, W.R., *et al*. (2003). Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. Nature *424*, 1042–1047.
103. Johnson, Z.I., Zinser, E.R., Coe, A., McNulty, N.P., Woodward, E.M., and Chisholm, S.W. (2006). Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. Science *311*, 1737–1740.
104. Strauss, S.Y., and Irwin, R.E. (2004). Ecological and evolutionary consequences of multispecies plant-animal interactions. Annu. Rev. Ecol. Evol. Systematics *35*, 435–466.
105. Weinbauer, M.G., and Rassoulzadegan, F. (2004). Are viruses driving microbial diversification and diversity? Environ. Microbiol. *6*, 1–11.
106. Riley, M.A., and Wertz, J.E. (2002). Bacteriocins: evolution, ecology, and application. Annu. Rev. Microbiol. *56*, 117–137.
107. Yim, G., Wang, H.H., and Davies, J. (2007). Antibiotics as signalling molecules. Philos. Trans. R. Soc. Lond. B Biol. Sci. *362*, 1195–1200.
108. Zhang, K., Martiny, A.C., Reppas, N.B., Barry, K.W., Malek, J., Chisholm, S.W., and Church, G.M. (2006). Sequencing genomes from single cells by polymerase cloning. Nat. Biotechnol. *24*, 680–686.
109. Du, J., Huang, Y.P., Xi, J., Cao, M.J., Ni, W.S., Chen, X., Zhu, J.K., Oliver, D.J., and Xiang, C.B. (2008). Functional gene-mining for salt tolerance genes with the power of Arabidopsis. Plant. J., epub ahead of print.