

**Wesleyan University**

---

**From the Selected Works of Frederick M. Cohan**

---

2009

# Tracking bacterial responses to global warming with an ecotype-based systematics

Frederick M Cohan, *Wesleyan University*



Available at: [https://works.bepress.com/frederick\\_cohan/10/](https://works.bepress.com/frederick_cohan/10/)

# Tracking bacterial responses to global warming with an ecotype-based systematics

F. M. Cohan

Department of Biology, Wesleyan University, Middletown, CT, USA

## Abstract

The broadly defined species of bacterial systematics frequently contain unnamed and unrecognized populations (ecotypes) differing in physiology, genome content, and ecology. Without formal recognition of such ecotypes, it is difficult for microbial ecologists to detect replacement of one ecotype by another in the face of global warming. The ecotype simulation algorithm has proved capable of supporting investigation of such replacements, as it has detected temperature-distinguished ecotypes that are invisible to the present bacterial systematics. Creating an ecotype-based systematics will help to identify the units of diversity that we will want to track as we seek to observe the early microbial responses to global warming.

**Keywords:** *Bacillus*, ecology, periodic selection, speciation, species concept

*Clin Microbiol Infect* 2009; **15** (Suppl.1): 54–59

**Corresponding author and reprint requests:** F. M. Cohan,  
Department of Biology, Wesleyan University, Middletown, CT  
06459–0170, USA  
E-mail: [fcohan@wesleyan.edu](mailto:fcohan@wesleyan.edu)

## Introduction

Thermometers and shrinking ice packs provide the physical alarms that the world's climate is warming, but biotic changes offer the most visceral harbingers of a hot future world—polar bears adrift on shrinking ice on an open Arctic Ocean, tropical birds appearing on the checklists of temperate birders, and European sheep succumbing to an African disease. As animals, plants and microbes move polewards, ecologists are challenged to track the responses of organisms to global warming. These efforts should help our species to accommodate to the biotic challenges of a warmer world. More generally, global warming presents a novel opportunity for microbial ecologists to investigate the roles of migration, adaptation and speciation in accommodating environmental changes.

Zoologists and botanists have already made significant progress in tracking the responses of many individual animal and plant species to global warming, e.g. showing how various marine fish have moved northwards in recent decades [1]. In addition, zoologists and botanists have predicted the future geographical responses of individual species. For example, the very closely related oak species *Quercus*

*douglasii* (blue oak) and *Quercus lobata* (valley oak) of California, with slightly different habitat requirements, have each been predicted to contract from warmer habitats and to expand into adjacent cooler habitats [2]. For our purposes, this case is particularly interesting, because the more drought-tolerant blue oak is invading the present habitat of the more mesophilic valley oak. Such a prediction is possible because, like most pairs of closely related plant and animal species, these oak species are narrowly defined so that each species is homogeneous within itself (at least at any one location), and distinct from the other in its physiology and preferred microhabitats [3,4]. Thus, the finely tuned systematics of plants (and animals) allows us to observe and predict the replacement of one extremely close relative by another in the face of global change.

We can imagine how difficult it would be to track or predict geographical range changes in oaks if plant systematics did not identify all of the closely related, ecologically distinct oak species within a region. Suppose instead that the only recognized oak taxon was the genus *Quercus* (i.e. with no individual species recognized), and that it was our job to track responses to global change in the genus *Quercus* at large. Changing our focus from the individual species to the *Quercus* amalgam, containing blue oak and valley oak, as well as dozens of other oak species from this region, would blind us to replacements of one subgroup by another. Unfortunately, this is exactly the situation that bacterial systematics leaves us in. The named, recognized species of bacterial systematics are defined quite broadly, much like a genus of animals or plants, such that the typical bacterial species is

extremely diverse in its physiology and genome content, but most fundamentally in its ecology [5,6].

The broad brush of bacterial systematics is seen clearly in the case of *Escherichia coli* [7]. Three strains within this species, one non-pathogenic, one uropathogenic, and the other enterohaemorrhagic, have been shown to share only 39% of their genes. Most significant, however, is the fact that these strains, so profoundly distinct in their ecology and clearly long divergent (having diverged so much in genome content), are judged by bacterial systematics to reside within the same species taxon. Given the wide vision of bacterial systematics for inclusion of diverse populations within a species, one can imagine that tracking or predicting the geographical responses of unnamed, unrecognized, ecologically distinct populations within a 'species' would be challenging, indeed.

We have recently proposed a paradigm shift in bacterial systematics that aims to incorporate ecological diversification into analysis of bacterial diversity [5,8]. Here, I will argue that this 'ecotype'-based systematics will enable microbiologists to observe geographical responses to global warming that would be invisible to the current systematics of bacteria. To this end, I will demonstrate that the new systematics can identify extremely closely related bacterial populations, with different temperature adaptations, that would be unrecognized by the prevailing systematics.

## The rationale for sequence-based demarcation of ecotypes

---

Identification of very closely related, ecologically distinct populations is more difficult in bacteriology than in zoology or botany [5]. Zoologists can anticipate, for a given animal group, the traits that determine the ecological niche. For example, a zoologist studying songbirds expects that the various species will differ in their bill shapes and sizes (which determine the kind and size of food consumed), and so can easily identify all the ecologically distinct populations within a community. However, a bacteriologist cannot anticipate with confidence the characteristics determining niche differences between closest relatives, even for a well-studied taxon. This is because the creation of new populations is frequently brought about through horizontal genetic transfer, where a new enzyme function or pathway is introduced into a recipient organism from any of a large set of potential donor organisms [5,9]. As we do not know what aspects of physiology to focus on, discovery of ecological diversity among closely related bacteria cannot depend on physiological analysis alone. Fortunately, DNA sequence surveys are well suited to discovering ecologically distinct bacterial populations ('ecotypes') [5].

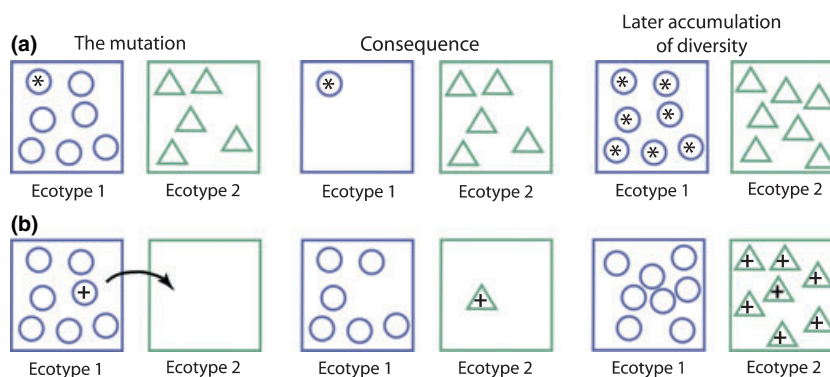
We have defined an ecotype as a clade of bacteria that are ecologically similar to one another, such that genetic diversity within the ecotype is limited by a cohesive force, either periodic selection or genetic drift, or both (Fig. 1) [5]. In this model, diversity within an ecotype is ephemeral, persisting only until the next periodic selection event, when diversity is brought to near zero at all loci, or until purged by genetic drift. Divergence becomes permanent when a mutation (or recombination event) places the organism into a new ecological niche and thereby founds a new ecotype. Because the new ecotype is ecologically distinct from the parental ecotype, periodic selection events in the parental ecotype cannot extinguish the founding organism and its descendants (Fig. 1). The new ecotype thus escapes the periodic selection events of the parental ecotype, and the two new ecotypes are free to diverge indefinitely. Ecotypes defined in this way bear the characteristics attributed to species by biologists outside of microbiology: each ecotype is cohesive (with diversity constrained by periodic selection and/or drift), different ecotypes may diverge indefinitely from one another, and the different ecotypes are ecologically distinct [10].

Such ecotypes may be revealed through molecular surveys, provided that one particular model of bacterial evolution holds. This is the stable ecotype model, where each ecotype persists through a long history of recurrent periodic selection events, and new ecotypes are only rarely formed. In this model, each ecotype is ideally expected to correspond to a separate DNA sequence cluster, for any gene shared among ecotypes, as periodic selection purges diversity at all loci, within but not between ecotypes, leading to a steady accrual of sequence divergence between ecotypes (Fig. 1).

## The ecotype simulation approach

---

Most sequence-based bacterial phylogenies are complex, with many levels of clusters and subclusters within clusters (Fig. 2), and so it is generally not clear which level of sequence cluster should correspond to ecotypes. Bacterial systematics has utilized various universal molecular criteria for demarcating clusters expected to be of biological significance. For example, species have been demarcated for decades under the guidance of a universal criterion of genome content similarity, as quantified by DNA-DNA hybridization [11]. More recently, species demarcation has been guided by divergence at the 16S rRNA locus, first with a 3% cut-off and more recently with a 1% cut-off [12]. However, there is no theoretical rationale for these cut-offs to correspond to biologically significant clades with species-like properties, and



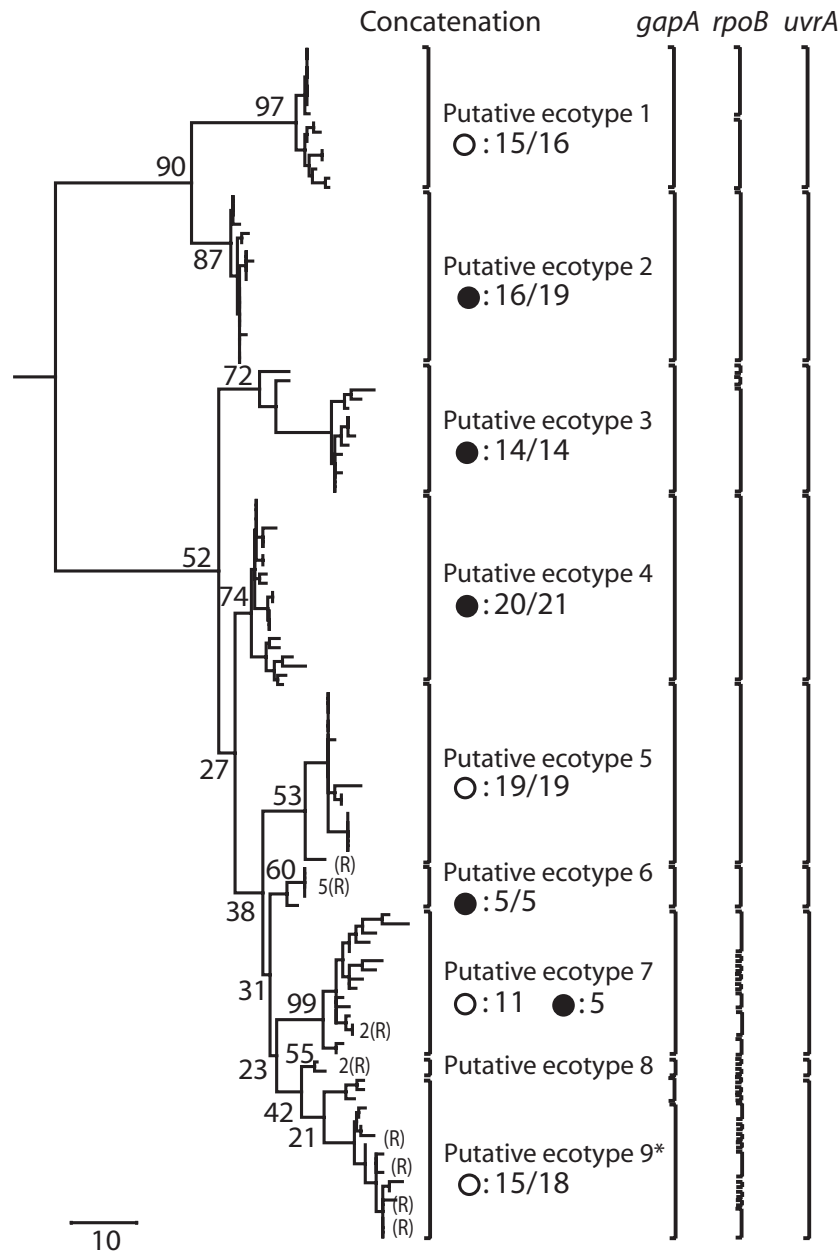
**Fig. 1.** Two classes of mutation and recombination events that determine ecotype diversity in bacteria. Circles and triangles represent different genotypes within ecotypes 1 and 2, respectively; asterisks represent adaptive mutations and crosses represent ecotype formation mutations. (a) Periodic selection mutations. These improve the fitness of an individual such that the mutant and its descendants (ecotype) outcompete all other cells within the ecological niche (i.e. cells of the same ecotype); these mutations do not affect the diversity within other ecotypes, because ecological differences prevent direct competition. Periodic selection leads to the distinctness of ecotypes by purging the divergence within but not between ecotypes. (b) Ecotype formation mutations. Here a mutation or recombination event allows the cell to occupy a new ecological niche, founding a new ecotype. The ecotype formation mutant, as well as its descendants, can no longer be extinguished by periodic selection events from its former ecotype [17]. (Used with permission from Landes Publishers.)

nor is it clear that any particular cut-off should apply to all bacteria [5]. In any case, applying the cut-offs embraced by systematists has led to the enormous ecological and physiological diversity seen within *E. coli* and within many other species.

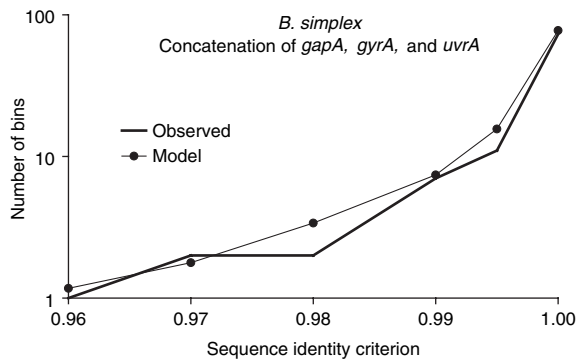
We have proposed a theory-based approach called ecotype simulation to derive cut-offs that are appropriate for demarcating a particular clade's ecotypes, allowing that different bacterial groups may have different cut-offs [8]. The ecotype simulation approach begins by characterizing the sequence diversity within a clade as the number of sequence clusters (or bins) present for different sequence identity criteria (Fig. 3) [13,14]. The number of sequence clusters at a particular sequence identity level represents the number of lineages at some point in the past that have survived to the present; thus, the sequence diversity curve represents the history of splitting of lineages within the clade [14]. The ecotype simulation algorithm estimates the rates of periodic selection and drift, the net rate of ecotype formation (taking into account ecotype extinction), and the number of ecotypes ( $n$ ), so as to yield a clade's sequence diversity pattern (Fig. 3) with maximum likelihood. Individual ecotypes are demarcated by determining the largest subclades that are each consistent with containing a single ecotype (i.e. such that  $n = 1$  for the subclade). Further details of ecotype simulation may be found in our previous work [8], and the software may be downloaded from <http://fcohan.web.wesleyan.edu/ecosim/>.

We have applied ecotype simulation to analyse 131 strains of *Bacillus simplex* isolated from two 'Evolution Canyons' of

Israel [8]. Each 'Evolution Canyon' is an arid canyon running east to west, providing two major habitats differing in solar insolation—these are the north-facing slope (NFS) and the south-facing slope (SFS). Ecotype simulation analysis inferred nine putative ecotypes within *B. simplex* (Fig. 2). We were able to confirm that many of these groups were ecologically distinct, first by comparing the putative ecotypes for their associations with the two major habitats. For example, at the top of Fig. 2 is a clade containing putative ecotypes 1 and 2, which were distinguished from one another by their strong associations with the SFS and NSF, respectively; other putative ecotypes in other clades were also distinguished by their associations with the NSF and SFS [8]. The ecological distinctness of the putative ecotypes was also corroborated by physiological differences. The SFS-associated ecotypes contained greater levels of high-temperature-adapting isomethyl-branched fatty acids than the NFS-associated ecotypes [15]. Also, the SFS ecotypes have shown higher growth rates than NFS ecotypes at a stressfully high temperature, whereas the differences disappeared at optimal temperatures. The NFS-adapted and SFS-adapted ecotypes were not different in their sensitivities to UV-C radiation, so temperature appears to be an important component of divergence between these ecotypes. We have not yet investigated other possible differences in niche among ecotypes, such as differences in organic resources utilized, or differences in interactions with other microbes. Other differences yet to be discovered may explain the coexistence of multiple ecotypes on the same slope, e.g. putative ecotypes 1, 5, 7 and 9 on the SFS.



**Fig. 2.** Phylogeny and ecotype demarcation of *Bacillus simplex* from the 'Evolution Canyons' of Israel. The phylogeny is based on a concatenation of three protein-coding genes, with recombinant sequences removed. The phylogeny contains many clusters and subclusters within clusters, and it is not intuitively clear how to demarcate the ecologically significant groups without a theory. Using ecotype simulation, ecotypes were demarcated as the most inclusive clades that were each consistent with being a single ecotype. Ecotype demarcations are indicated by brackets, as based on analysis of the concatenation as well as each individual gene. The ecotype demarcations were similar as based on the concatenation and the individual genes, except that the more rapidly evolving gene *rpoB* tended to split the ecotypes determined by analysis of the concatenation. A group of related recombinants is indicated by 'R' following the number of recombinants. For isolates that had recombined at one gene locus, ecotype placement was determined by ecotype simulation of the two genes that had not recombined. With one exception, demarcated ecotypes were supported as monophyletic groups in at least 50% of bootstrap replications (percentage bootstrap support indicated at nodes); the exception is asterisked to indicate that its phylogenetic status is tentative, pending additional sequence data. Microhabitat sources were the south-facing slope (○) and the north-facing slope (●). For each ecotype represented by at least four isolates, the principal microhabitat source(s) is indicated. If one microhabitat provided at least 80% of the isolates, the principal microhabitat source is indicated; for ecotypes not so dominated by a single source, all microhabitat sources are indicated. Note that the prevailing practice of bacterial systematics has included all of this diversity within one species [8]. (Used with permission from the National Academy of Sciences.)



**Fig. 3.** Observed and modelled clade sequence diversity patterns. Sequences for a gene (or a concatenation of genes) were binned into clusters with different levels of minimum pairwise identity. The curves represent the diversity among 116 *Bacillus simplex* isolates from 'Evolution Canyons' I and II, as based on a concatenation of *gapA*, *rpoB*, and *uvrA*, with 15 recombinant organisms removed [8]. The individual points for the model curve are means based on 1000 replications of the maximum-likelihood solution. (Used with permission from the National Academy of Sciences.)

All of the *B. simplex* ecotypes from the 'Evolution Canyons' are extremely closely related to one another, as revealed by complete identity in their 16S rRNA sequences [8]. Thus, in all likelihood, these ecotypes would remain unrecognized and unnamed with the prevailing practice of bacterial systematics [8].

We have also applied ecotype simulation to a clade of hot spring cyanobacteria from Yellowstone National Park in the USA [16]. Within the A and A' subclades of *Synechococcus*, ecotype simulation identified putative ecotypes that were confirmed to be ecologically distinct by differences in associations with temperature and depth in the photic zone. Some of these confirmed ecotypes were only 0.7% divergent at 16S rRNA, and so would probably be unrecognized by bacterial systematics.

## Towards an ecotype-based systematics

Ecotype simulation promises to identify the extremely closely related ecotypes that have up to now been included within the recognized species of bacterial systematics. We have previously proposed a protocol for incorporating ecotypes and ecological diversification into bacterial systematics, taking into account that factors other than periodic selection may contribute to sequence clustering in certain lineages [5,8]. The first step is to infer putative ecotypes through ecotype simulation analysis of DNA sequence data (or through another theory-based model of bacterial evolution and

ecology). The second step is to confirm that the putative ecotypes so identified are actually distinct in their ecology in nature. This could involve comparison of microgeographical distribution, physiology, genome content and/or genome-wide gene expression among putative ecotypes. We have suggested that ecotypes discovered within the phylogenetic range of an existing, named species (e.g. within 1% divergence at 16S rRNA) should be named as a trinomial 'ecovar' within the established species; also, newly discovered ecotypes that are outside the phylogenetic range of existing species should each be named as a separate species [5].

Recognition of ecotypes will yield the systematic infrastructure with which to discern early and subtle responses to global warming. Just as botanists can predict or track expansion of the drought-tolerant blue oak into the cooler habitats now held by the valley oak, the ecotype-based systematics will allow microbial ecologists to track replacement of a mesophilic ecotype by an extremely close relative adapted to hotter microclimates. Ecotype simulation has proved capable of supporting such observations in the future, as it has detected temperature-distinguished ecotypes that are invisible to the present systematics, in the cases of *Bacillus* and *Synechococcus* [8,16]. I expect that many broadly defined species recognized by systematics will be shown to contain temperature-distinguished ecotypes as well. These are the units of diversity that we will want to track as we seek to observe the early microbial responses to global warming.

## Transparency Declaration

This work was funded by NSF FIBR award EF-0328698 and by grants from Wesleyan University. The author received an honourarium from the Lilly Foundation for speaking at their symposium on global change.

## References

- Perry AL, Low PJ, Ellis JR, Reynolds JD. Climate change and distribution shifts in marine fishes. *Science* 2005; 308: 1912–1915.
- Kueppers LM, Snyder MA, Sloan LC, Zavaleta ES, Fulfrost B. Modeled regional climate change and California endemic oak ranges. *Proc Natl Acad Sci USA* 2005; 102: 16281–16286.
- Grant V. *Plant speciation*. New York: Columbia University, 1971.
- Mayr E. *The growth of biological thought: diversity, evolution, and inheritance*. Cambridge: Harvard University Press, 1982.
- Cohan FM, Perry EB. A systematics for discovering the fundamental units of bacterial diversity. *Curr Biol* 2007; 17: R373–R386.
- Staley JT. The bacterial species dilemma and the genomic-phylogenetic species concept. *Phil Trans R Soc Lond B Biol Sci* 2006; 361: 1899–1909.

7. Welch RA, Burland V, Plunkett G III et al. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* 2002; 99: 17020–17024.
8. Koeppl A, Perry EB, Sikorski J et al. Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc Natl Acad Sci USA* 2008; 105: 2504–2509.
9. Gogarten JP, Doolittle WF, Lawrence JG. Prokaryotic evolution in light of gene transfer. *Mol Biol Evol* 2002; 19: 2226–2238.
10. de Queiroz K. Ernst Mayr and the modern concept of species. *Proc Natl Acad Sci USA* 2005; 102 (suppl 1): 6600–6607.
11. Wayne LG, Brenner DJ, Colwell RR et al. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* 1987; 37: 463–464.
12. Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006; 33: 152–155.
13. Acinas SG, Klepac-Ceraj V, Hunt DE et al. Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 2004; 430: 551–554.
14. Martin AP. Phylogenetic approaches for describing and comparing the diversity of microbial communities. *Appl Environ Microbiol* 2002; 68: 3673–3682.
15. Sikorski J, Nevo E. Patterns of thermal adaptation of *Bacillus simplex* to the microclimatically contrasting slopes of ‘Evolution Canyons’ I and II, Israel. *Environ Microbiol* 2007; 9: 716–726.
16. Ward DM, Bateson MM, Ferris MJ et al. 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 R Soc Ser B* 2006; 361:1997–2008.
17. Cohan FM. Periodic selection and ecological diversity in bacteria. In: Nurminsky D, ed. *Selective sweep*. Georgetown, TX: Landes Bioscience, 2005; 78–93.