

The importance of resolving biogeographic patterns of microbial microdiversity



Alexander B Chase

Department of Ecology
and Evolutionary Biology
University of California
Irvine, CA, USA
Email: abchase@uci.edu



Jennifer BH Martiny

Department of Ecology
and Evolutionary Biology
University of California
Irvine, CA, USA
Email: jmartiny@uci.edu

For centuries, ecologists have used biogeographic patterns to test the processes governing the assembly and maintenance of plant and animal communities. Similarly, evolutionary biologists have used historical biogeography (e.g. phylogeography) to understand the importance of geological events as barriers to dispersal that shape species distributions. As the field of microbial biogeography initially developed, the utilisation of highly conserved marker genes, such as the 16S ribosomal RNA gene, stimulated investigations into the biogeographic patterns of the microbial community as a whole. Here, we propose that we should now consider the biogeographic patterns of microdiversity, the fine-scale genetic diversity observed within a traditional ribosomal-based operational taxonomic unit.

Biogeography investigates how ecological and evolutionary processes influence the distribution of biodiversity and the structure of contemporary communities¹. Historically, biogeographic patterns of plants and animals are studied at the species level and describe large-scale patterns of species' distributions. In contrast, the vast majority of microbial biogeographic studies investigate patterns by sampling the entire community at broad taxonomic designations. Typically, these studies define operational taxonomic units (OTUs) using a highly conserved ribosomal marker gene, usually the 16S rRNA gene for bacteria and archaea. However, the decision of which genetic region to target, and in particular the genetic resolution of that region, can influence the biogeographic patterns observed². While these conserved regions can capture a large breadth of the microbial community, these regions, by their very nature, limit the detection of finer-scale genetic variation. By resolving diversity within the OTU designation, we can detect ecological and

evolutionary processes occurring at this fine taxonomic scale that might otherwise be overlooked.

What OTU-based biogeography can and can't tell us

It is now well established that microbial communities assayed by traditional OTU designations display distinct biogeographic patterns over space and time. These patterns have been identified in environments ranging from marine³, to terrestrial⁴, and to human-associated systems⁵. Combined with abiotic and biotic data from the sampled environment, such patterns can provide initial hypotheses about the ecological processes shaping microbial community assemblages⁶. Thousands of microbial studies now demonstrate that OTU-based patterns primarily reflect the importance of selection of environmental conditions based on correlations between microbial composition and the environment (Figure 1a). These patterns indicate that OTUs comprising each microbial community vary in their ability to tolerate various abiotic and biotic conditions, suggesting partitioning of environmental resources and niche spaces among taxa in the community.

While environmental variables explain much of the variation in microbial composition, many studies also find that some variation is correlated to the geographic distances between communities⁶. This observation can be illustrated with a distance-decay curve, or a negative correlation between the similarity in microbial composition with geographic distance between pairwise samples⁷ (Figure 1b). If this negative relationship holds after accounting for environmental variation, then the pattern suggests that ecological drift, caused by stochastic fluctuations in demographic patterns, contributes to variation in community composition^{8,9}. Further,

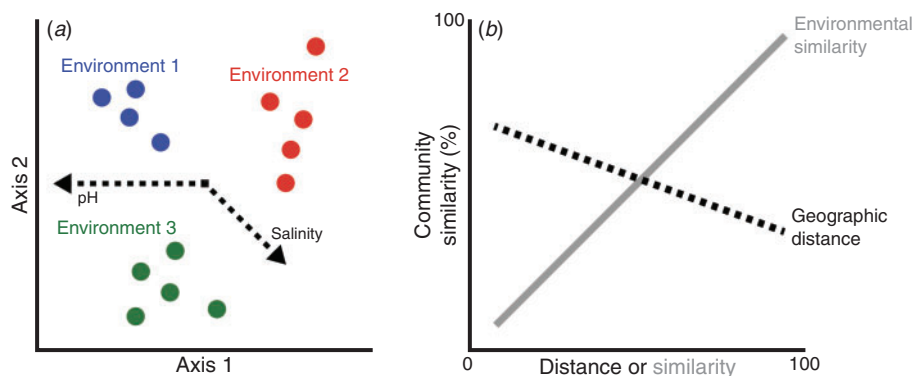


Figure 1. Hypothetical community analyses from OTU-based studies. (a) An ordination plot depicting community composition across three environments with the main environmental factors driving compositional differences indicated with dashed arrows. Each point represents a sampled microbial community, with points closer to one another indicating higher similarity in community composition. (b) Community similarity among a collection of samples is often positively correlated to environmental similarity (grey line) and negatively correlated with geographic distance (black dashed line, also called a distance-decay curve). The influence of strong environmental selection on the community is reflected in the positive correlation with increasing environmental similarity, while the influence of ecological drift is reflected in the negative correlation with increasing geographic distance between samples.

since ecological drift depends on restricted dispersal, the pattern gives insight into the degree of dispersal limitation between the sampled communities. A caveat to such studies is that it is impossible to completely account for environmental variation, and the environment is spatially autocorrelated. However, such OTU-based studies suggest that the ecological processes of both environmental selection and ecological drift contribute to biogeographic patterns at this broad genetic resolution⁷.

In contrast to ecological processes, biogeographic patterns of OTU-based analyses are unlikely to detect patterns shaped by evolutionary processes. This limitation is due to the broad resolution of conserved marker genes. Variation in these genetic regions capture relatively distant evolutionary divergences, especially when clustered at 97% sequence similarity. Indeed, a 3% sequence divergence in the 16S rRNA gene, the most common level of OTU clustering, represents roughly 150 million years of evolutionary history¹⁰, or before the origin of modern birds¹¹. In other words, biogeographic patterns for birds at this taxonomic level would mask all diversification within the group. Similarly, the use of such conserved marker genes for microbes will generally preclude detecting biogeographic patterns emerging from evolutionary processes, such as endemism and niche conservatism, as observed for macroorganisms assessed at the species or population level.

What is microbial microdiversity

Studies based on 16S rRNA sequences have been instrumental in identifying ecological patterns and their underlying processes at relatively broad genetic resolutions. However, it is increasingly clear that there is extensive genetic diversity within 16S-based OTUs, so-called microdiversity, in environmental habitats^{12,13}. For example, a natural population of the bacterioplankton *Vibrio splendidus* contained >1000 distinct genotypes, even when

clustered at >99% 16S rRNA sequence similarity¹⁴. Based on their very nature, conserved marker genes lack the variability to resolve fine-scale diversity within an OTU. Even with the implementation of exact sequence variants (ESVs), the 16S rRNA gene simply cannot resolve the fine-scale variation among closely related microbial lineages¹⁵. Thus, different approaches are needed to investigate the biogeographic patterns of this vast genetic diversity.

Beyond identifying genetic microdiversity, a key question is whether this genetic variation is phenotypically relevant¹⁶. Investigations into microdiverse marine bacterial taxa suggest that they vary in physiological traits including preferences for particular abiotic conditions^{13,17}. Further, some of this trait variation within OTU-based taxa appears to be phylogenetically conserved within microdiverse clades^{18,19}, although resolving the phylogeny of such closely related strains is often difficult with 16S rRNA sequences (Figure 2a). Instead, taxon-specific marker genes or, ideally full genome sequences, can often resolve microdiverse clades and reveal which traits are shared among particular phylogenetic clades (Figure 2b). For example, an analysis of strain diversity of an abundant leaf litter bacterium, *Curtobacterium*, exhibited extensive variation in the degree of polymeric carbohydrate degradation and temperature preference among microdiverse clades²⁰. Thus, more resolved genetic and physiological studies can help to establish the phylogenetic distribution of traits.

What biogeographic patterns of microdiversity can tell us

The presence of trait variation among microdiverse clades suggests that microdiversity will exhibit distinctive biogeographic patterns. If this trait variation corresponds to different ecological preferences, then the environment should select for specific clades under variable conditions. Indeed, different bacterial ecotypes,

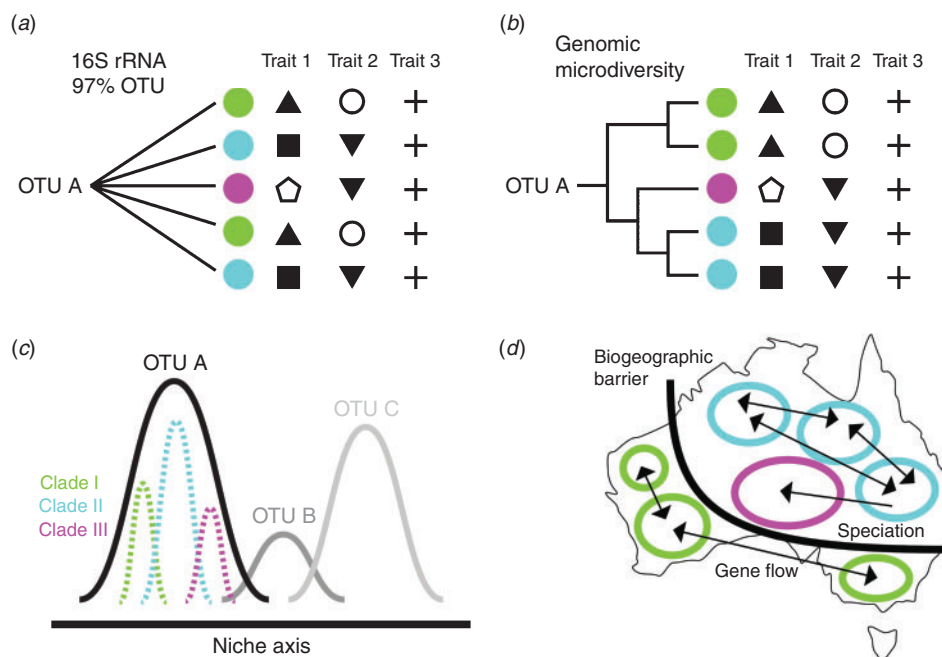


Figure 2. Detection of ecological and evolutionary processes within OTU A with microdiverse Clades I (green), II (blue), and III (pink). (a) The 16S rRNA gene often cannot resolve phylogenetic relationships within a 16S-based OTU and, subsequently, the distribution of traits among clades. (b) Genomic sequences or multi-locus sequence analyses (MLSA) of marker genes can resolve phylogenetic relationships at a finer-scale revealing, in this hypothetical example, that strains within clades share more similar traits. (c) Trait variation within microdiverse taxa can promote resource partitioning in the environment leading to fine-scale niche differentiation among clades (represented in colored dashed lines) that would otherwise be masked at the OTU level (black line represents the total niche for OTU A). (d) Investigating genetic differentiation within OTUs is more likely to reveal dispersal limitation (measured by gene flow between clade populations) and the presence of biogeographic barriers that contribute to microbial diversification. In this hypothetical example, black arrows represent gene flow between populations of microdiverse clades, where limited gene flow (no arrows connecting green with the blue and pink populations) suggests the presence of biogeographic barriers.

or ecological populations, have repeatedly been shown to vary in their spatial distribution. Thus, closely-related clades appear to partition niche space in the environment that would normally be masked at the OTU level (Figure 2c). For example, at the OTU level, the globally distributed cyanobacterium, *Prochlorococcus*, shows a broad preference for low-nutrient and warmer waters²¹. However, microdiverse clades of *Prochlorococcus* exhibit distinct spatial distribution patterns shaped by additional environmental factors, including light availability and temperature^{12,17,22}. Thus, biogeographic patterns of microdiversity can elucidate the importance of key environmental parameters governing niche differentiation that may not be identifiable at the OTU designation.

Perhaps even more importantly, a focus on microdiversity can reveal evolutionary processes that would otherwise be masked at a broader genetic resolution. Thus far, few environmental studies have targeted microbial diversity at a fine enough scale to investigate how evolutionary mechanisms, such as mutation and genetic drift, can lead to differential biogeographic patterns^{18,23}. Those examples that do exist find evidence for evolutionary processes contributing to spatial patterns. In one such example, reduced dispersal between hot spring populations of the archaeon thermophile *Sulfolobus*, restricted gene flow to allow diversification to occur among geographic regions^{24,25}. Similarly in terrestrial soils, dispersal limitation at regional spatial scales structures bacterial

populations of *Streptomyces* along a latitudinal gradient²⁶. With the increased availability of computational tools to study population genomics²⁷ and the incorporation of gene flow networks²⁸, we expect that more studies will consider the spatial distribution of microdiversity. Such studies are likely to illuminate the effects of evolutionary processes on microbial diversity in the environment, including the presence of biogeographic barriers and the degree of microbial endemism²⁹ (Figure 2d).

Conclusions

Future progress in microbial biogeography necessitates moving beyond the OTU designation. While OTU-based studies will continue to play an important role in microbial biogeography, an intensified focus on finer-genetic diversity will uncover thus-far unidentified ecological and evolutionary patterns. However, these studies will require targeted sampling of particular microbial taxa rather than the entire community. Generally, this effort will require moving beyond targeting the 16S rRNA gene; even ESVs of this region will not be able to distinguish microbial populations at a fine enough genetic scale. And while extensive shotgun metagenomic and targeted amplicon sampling can reveal co-occurrence of novel microdiversity associated with distinct environmental conditions³⁰, these studies are dependent on the interpretation of genomic potential for ecological diversity. Therefore, there is still a need

to link the genomic variation to functional traits that will define ecotypes. The return to isolation-based studies to gather relevant genetic and physiological information will better inform environmental metagenomic studies investigating microbial microdiversity. By expanding the focus to microbial microdiversity and implementing targeted environmental studies, we can better understand the ecological and evolutionary processes generating microbial biogeographic patterns as macroecologists have done for decades.

References

- Wiens, J.J. and Donoghue, M.J. (2004) Historical biogeography, ecology and species richness. *Trends Ecol. Evol.* **19**, 639–644. doi:10.1016/j.tree.2004.09.011
- Cho, J.-C. and Tiedje, J.M. (2000) Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Appl. Environ. Microbiol.* **66**, 5448–5456. doi:10.1128/AEM.66.12.5448-5456.2000
- Giovannoni, S.J. *et al.* (1996) 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the green non-sulfur bacteria. *Proc. Natl. Acad. Sci. USA* **93**, 7979–7984. doi:10.1073/pnas.93.15.7979
- Fierer, N. and Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. USA* **103**, 626–631. doi:10.1073/pnas.0507535103
- Human Microbiome Project Consortium. (2012) Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214. doi:10.1038/nature11234
- Hanson, C.A. *et al.* (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* **10**, 497–506. doi:10.1038/nrmicro2795
- Martiny, J.B.H. *et al.* (2006) Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* **4**, 102–112. doi:10.1038/nrmicro1341
- Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.
- McGill, B.J. *et al.* (2006) Empirical evaluation of neutral theory. *Ecology* **87**, 1411–1423. doi:10.1890/0012-9658(2006)87[1411:EEONT]2.0.CO;2
- Ochman, H. *et al.* (1999) Calibrating bacterial evolution. *Proc. Natl. Acad. Sci. USA* **96**, 12638–12643. doi:10.1073/pnas.96.22.12638
- Pereira, S.L. and Baker, A.J. (2006) A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. *Mol. Biol. Evol.* **23**, 1731–1740. doi:10.1093/molbev/msl038
- Moore, L.R. *et al.* (1998) Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* **393**, 464–467. doi:10.1038/30965
- Jaspers, E. and Overmann, J. (2004) Ecological significance of microdiversity: identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysologies. *Appl. Environ. Microbiol.* **70**, 4831–4839. doi:10.1128/AEM.70.8.4831-4839.2004
- Thompson, J.R. *et al.* (2005) Genotypic diversity within a natural coastal bacterioplankton population. *Science* **307**, 1311–1313. doi:10.1126/science.1106028
- Lan, Y. *et al.* (2016) Marker genes that are less conserved in their sequences are useful for predicting genome-wide similarity levels between closely related prokaryotic strains. *Microbiome* **4**, 18. doi:10.1186/s40168-016-0162-5
- Larkin, A.A. and Martiny, A.C. (2017) Microdiversity shapes the traits, niche space, and biogeography of microbial taxa. *Environ. Microbiol. Rep.* **9**, 55–70. doi:10.1111/1758-2229.12523
- Johnson, Z.I. *et al.* (2006) Partitioning among *Prochlorococcus* ecotypes along environmental gradients. *Science* **311**, 1737–1740. doi:10.1126/science.1118052
- Martiny, A.C. *et al.* (2009) Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. *Environ. Microbiol.* **11**, 823–832. doi:10.1111/j.1462-2920.2008.01803.x
- Martiny, J.B.H. *et al.* (2015) Microbiomes in light of traits: a phylogenetic perspective. *Science* **350**, aac9323. doi:10.1126/science.aac9323
- Chase, A.B. *et al.* (2017) Microdiversity of an abundant terrestrial bacterium encompasses extensive variation in ecologically relevant traits. *MBio* **8**, e01809-17. doi:10.1128/mBio.01809-17
- Flombaum, P. *et al.* (2013) Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci. USA* **110**, 9824–9829. doi:10.1073/pnas.1307701110
- Martiny, A.C. *et al.* (2006) Phosphate acquisition genes in *Prochlorococcus* ecotypes: evidence for genome-wide adaptation. *Proc. Natl. Acad. Sci. USA* **103**, 12552–12557. doi:10.1073/pnas.0601301103
- Andam, C.P. *et al.* (2016) A latitudinal diversity gradient in terrestrial bacteria of the genus *Streptomyces*. *MBio* **7**, e02200-15. doi:10.1128/mBio.02200-15
- Whitaker, R.J. *et al.* (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* **301**, 976–978. doi:10.1126/science.1086909
- Cadillo-Quiroz, H. *et al.* (2012) Patterns of gene flow define species of thermophilic Archaea. *PLoS Biol.* **10**, e1001265. doi:10.1371/journal.pbio.1001265
- Choudoir, M.J. *et al.* (2016) Latitude delineates patterns of biogeography in terrestrial *Streptomyces*. *Environ. Microbiol.* **18**, 4931–4945. doi:10.1111/1462-2920.13420
- Shapiro, B.J. *et al.* (2012) Population genomics of early events in the ecological differentiation of bacteria. *Science* **336**, 48–51. doi:10.1126/science.1218198
- Hehemann, J.-H. *et al.* (2016) Adaptive radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes. *Nat. Commun.* **7**, 12860. doi:10.1038/ncomms12860
- Polz, M.F. *et al.* (2013) Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends Genet.* **29**, 170–175. doi:10.1016/j.tig.2012.12.006
- Malmstrom, R.R. *et al.* (2013) Ecology of uncultured *Prochlorococcus* clades revealed through single-cell genomics and biogeographic analysis. *ISME J.* **7**, 184–198. doi:10.1038/ismej.2012.89

Biographies

Alexander B Chase is a doctoral student at the University of California, Irvine under Dr Jennifer Martiny. His research interests include applying his previous training in ecological theory, where he worked under Dr Stephen Hubbell at the University of California, Los Angeles, to microbial systems. He utilises bioinformatics to wrangle genomic and metagenomic data to investigate the ecological and evolutionary processes affecting fine-scale diversity in environmental bacterial taxa.

Jennifer BH Martiny is a Professor of Ecology and Evolutionary Biology and Director of the Microbiome Initiative at the University of California, Irvine. She is a fellow of the Ecological Society of America and the American Academy of Microbiology. Her lab's research aims to uncover fundamental principles of the generation and maintenance of diversity in microbial communities.

Corrigendum

In *Microbiology Australia* (Volume 38, Issue 4, pp. 204–205), the organisation name 'World Federation of Culture Collections' should be 'World Federation for Culture Collections' throughout.