



# Multi-omic Characterization of Intraspecies Variation in Laboratory and Natural Environments

Megan G. Behringer<sup>a,b,c</sup>

<sup>a</sup>Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, USA

<sup>b</sup>Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>c</sup>Vanderbilt Microbiome Initiative, Nashville, Tennessee, USA

**ABSTRACT** Investigation of microbial communities has led to many advances in our understanding of ecosystem function, whether that ecosystem is a subglacial lake or the human gut. Within these communities, much emphasis has been placed on inter-specific variation and between-species relationships. However, with current advances in sequencing technology resulting in both the reduction in sequencing costs and the rise of shotgun metagenomic sequencing, the importance of intraspecific variation and within-species relationships is becoming realized. Our group conducts multi-omic analyses to understand how spatial structure and resource availability influence diversification within a species and the potential for long-term coexistence of multiple ecotypes within a microbial community. Here, we present examples of ecotypic variation observed in the lab and in the wild, current challenges faced when investigating intra-species diversity, and future developments that we expect to define the field over the next 5 years.

**KEYWORDS** ecotypic variation, experimental evolution, population genomics, systems biology, metagenomics, systems microbiology

Estimated to account for almost five million species, prokaryotes are the most diverse and abundant organisms on earth (1). Prior to the genomic era, microbiologists suspected this incredible diversity in microbial species, but the lack of a universal marker hindered the ability to effectively describe the vast spectrum of microbial life. Development of phylogenetic approaches based on the 16S ribosomal DNA loci by Carl Woese et al. not only resolved *Archaea* as a distinct domain of the prokaryotic tree of life but also provided a way to distinguish divergent prokaryotic taxa and assess evolutionary relatedness (2, 3). Yet, despite the descriptive capability of this universal locus, investigation of microbial communities was still largely culture dependent until the late 2000s, and a swath of diversity represented by “unculturable” bacteria remained elusive. In response, efforts to harness culture-independent techniques were made and protocols utilizing microarrays (4) and pyrosequencing (5) were subsequently developed. Ultimately, it was the pairing of 16S rRNA gene analysis with next-generation sequencing which resulted in a widely accessible, low-cost, and high-throughput technique for the broad quantification of microbial diversity. In the 10 years since, 16S amplicon sequencing has fueled thousands of microbiome studies (6). These investigations have been vital to describing the tremendous role of microbes in maintaining ecosystem function and suggest that the composition of microbial communities can act as an indicator of host or environmental ecosystem health (7–9). While 16S sequencing has been powerful for resolving microbial diversity at the level of genera and species, considerable advances in computational power and sequencing technology have further revealed a wealth of intraspecies genetic variation through shotgun metagenomic sequencing. As such,

**Citation** Behringer MG. 2021. Multi-omic characterization of intraspecies variation in laboratory and natural environments. *mSystems* 6:e00764-21. <https://doi.org/10.1128/mSystems.00764-21>.

**Copyright** © 2021 Behringer. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to [megan.g.behringer@vanderbilt.edu](mailto:megan.g.behringer@vanderbilt.edu).

Conflict of Interest Disclosures: M.G.B. has nothing to disclose.

The views expressed in this article do not necessarily reflect the views of the journal or of ASM.

This article is part of a special series sponsored by [Floré](https://www.floridianscience.com/).

**Published** 24 August 2021

the functional diversity of microbial communities may extend beyond the species level, and the presence of multiple strains or ecotypes within a microbial species population can also be functionally important (10).

Subpopulation structure and ecotypic diversification is a common evolutionary outcome in heterogeneous environments. In the lab, populations of *Pseudomonas fluorescens* and *Escherichia coli* were demonstrated to diversify spatially based on varying ability to form biofilms and preference for the surface-air interface niche (11, 12). When specifically selected to progress through regular biofilm cycles due to cultivation in batch culture with plastic beads, *Burkholderia cenocepacia* populations were observed to diversify into three ecotypes representing a network of cross-feeding interactions (13). Fixing the spatial location of *E. coli* cells by culture in microfluidic channels similarly resulted in the creation of metabolic microhabitats. Here, ecotypes arose based on cross-feeding between glucose-fermenting cells and cells that evolved to utilize acetate waste products as a source of essential carbon.

While the nature of heterogeneous environments likely encourages diversification due to the presence of structured niches, adaptive diversification is also observed in more homogeneous conditions. Richard Lenski's Long-term Evolution Experiment (LTEE) bacteria, a collection of 12 *E. coli* populations that have been famously evolving for over 70,000 generations, have been cultivated in flasks. Despite the LTEE's relatively simple and well-mixed culture environment, more than half of the populations have exhibited the evolution of long-term coexisting ecotypes (14, 15), including a population where diversification resulted in the novel ability to aerobically uptake and utilize citrate (16). Another simple environment supporting the evolution of ecotypes is continuous culture in chemostats, where again the emergence of glucose and acetate specialists was observed in *E. coli* populations (17, 18).

While it would be easy to dismiss these examples of diversification as a laboratory artifact and a result of monospecies populations cultivated in the absence of interspecific competitors, significant within-species variation is also found in clinical and environmental settings. For example, in cystic fibrosis lung-like conditions, *Pseudomonas aeruginosa* was reported to rapidly diversify (19), and long-term coexistence of ecotypes has been observed in clinical samples from a single host (20). Moreover, sampling of coastal waters revealed coexisting ecotypes of the bacterioplankton *Vibrio splendidus* that were likely based on resource partitioning (21), and *Curtobacterium* soil ecotypes exhibit specialization along a climate gradient (22). These examples illustrate that ecotypic variation is likely widespread across environmental contexts. Thus, there is a need for more in-depth investigation of how microbes adapt and how different genetic backgrounds or culture conditions permit/inhibit ecotypic diversification.

Our group's research is interested in how interspecies variation is generated and maintained in bacterial populations. These studies can range from investigations of spontaneous mutation rates using a mutation accumulation/whole-genome sequencing (MA/WGS) approach to experimental evolution studies examining how space and stress shape bacterial adaptation. Using a multi-omic approach that combined time-series metagenomic sequencing with genomic, transcriptomic, and exometabolomic analysis of isolated clones, we demonstrated that daily-transferred *E. coli* populations maintained in spatially complex (16- by 100-mm glass culture tubes) and nutrient-rich (LB-Miller broth) media rapidly diversify into ecotypes based on location (biofilm versus planktonic) and then utilization of amino acid resources (12). These ecotypes coexisted long-term (>10,000 generations) and were a specific adaptive outcome of culture under our heterogeneous conditions, as their evolved fitness benefits did not translate to more-homogeneous environments.

Going forward, our group will focus more on how environment influences the evolution of ecotypes and if any genetic backgrounds are particularly permissive/inhibitory to the maintenance of subpopulation structure (23). Beyond lab-based studies, we

also intend to apply our understanding of adaptive diversification and bacterial population genetics to clinical investigations. As has been established, if bacteria are given the niche availability, then they can diversify and inhabit novel spaces. This potential to evolve subpopulations also represents the ability to maintain multiple genetic backgrounds on which antimicrobial resistance can emerge. Thus, understanding how pathogen populations evolve throughout the course of an infection is of major interest to us, especially in recurrent contexts. To this end, we are partnering with physicians, clinical microbiologists, and microbiologists specializing in pathogenesis to gain a deeper understanding of microbial population dynamics and their effect on disease.

## CHALLENGES AND FUTURE OF THE FIELD

To date, the identification of ecotypes and the characterization of their interactions remain challenging. On the genomic level, short reads produced by next-generation sequencing platforms provide the accuracy to confidently identify polymorphisms within a bacterial population but fall short in the ability to phase these polymorphisms across sequencing reads to determine which mutations are present within an individual. To overcome this issue, computational approaches have been developed to assist in the identification of ecotypes, such as clonal reconstruction (a clone-informed maximum-likelihood method to parsing time-series metagenomic sequencing of populations [24]) or clade-aware HMM (a hidden Markov model that infers if a mutation is fixed or polymorphic within a basal, major, or minor clade from time-series sequencing [15]). While these methods have been proven powerful for resolving subpopulation structure when applied to experimental evolution studies, they are not without their shortcomings: clonal reconstruction relies on additional sequencing of clonal isolates, which is problematic for unculturable taxa, clade-aware HMM struggles with more complex population structures, and both are computationally challenging. Future development of these and other computational approaches implementing machine learning, as well as current progress increasing the accuracy of single-bacterial-cell DNA sequencing (25) and single-molecule sequencing platforms, will increase the accessibility of studying the diversity within microbial populations in a broad variety of contexts.

Once ecotypes are identified, characterizing their role within the population or microbiome has additional complications. Microbes often exhibit emergent behavior, and their phenotypic expression may significantly vary in isolation from how they respond in the presence of their community (26). On a broader scale, this behavior can be assessed with metatranscriptomics. However, due to challenges introduced by genetic relatedness, assessing ecotype-specific differences in global gene expression from coculture is impossible, since most sequencing reads will lack any identifying information that could otherwise indicate from which ecotype that transcript may have originated. Fortunately, current developments in bacterial single-cell RNA sequencing methods (27–29) which bypass the complications of extremely low RNA yields from prokaryotic cells and nonpolyadenylated mRNAs will heavily mitigate this issue in the future. These methods will not only permit the characterization of ecotypes but also distinguish differential gene expression for closely related species in coculture and investigations of phenotypic heterogeneity (30). Lastly, without time-series sampling, true subpopulation structure cannot be distinguished from transient polymorphism on its way to fixation or extinction. As such, biobanking of microbiome samples will become increasingly important to investigate within-species variation in clinical or environmental contexts. Connecting the diversity within these communities to the presence of pathogenic or ecologically important isolates will empower microbiologists to identify microbiomes in transition between healthy and disease states so that future interventions may be developed.

## ACKNOWLEDGMENTS

I thank B.J.H., C.J.S., and S.R.B. for their helpful comments.

Our work investigating environmental effects on intraspecies variation is currently funded by the Army Research Office award W911NF-21-1-0161.

## REFERENCES

1. Louca S, Mazel F, Doebeli M, Parfrey LW. 2019. A census-based estimate of Earth's bacterial and archaeal diversity. *PLoS Biol* 17:e3000106. <https://doi.org/10.1371/journal.pbio.3000106>.
2. Woese CR, Fox GE. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* 74:5088–5090. <https://doi.org/10.1073/pnas.74.11.5088>.
3. Woese CR, Fox GE, Zablen L, Uchida T, Bonen L, Pechman K, Lewis BJ, Stahl D. 1975. Conservation of primary structure in 16S ribosomal RNA. *Nature* 254:83–86. <https://doi.org/10.1038/254083a0>.
4. Brodie EL, Desantis TZ, Joyner DC, Baek SM, Larsen JT, Andersen GL, Hazen TC, Richardson PM, Herman DJ, Tokunaga TK, Wan JM, Firestone MK. 2006. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Appl Environ Microbiol* 72:6288–6298. <https://doi.org/10.1128/AEM.00246-06>.
5. Roesch LFW, Fulthorpe RR, Riva A, Casella G, Hadwin AKM, Kent AD, Daroub SH, Camargo FAO, Farmerie WG, Triplett EW. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* 1:283–290. <https://doi.org/10.1038/ismej.2007.53>.
6. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* 108(Suppl 1):4516–4522. <https://doi.org/10.1073/pnas.100080107>.
7. Eisenstein M. 2020. The hunt for a healthy microbiome. *Nature* 577:S6–S8. <https://doi.org/10.1038/d41586-020-00193-3>.
8. Manor O, Dai CL, Kornilov SA, Smith B, Price ND, Lovejoy JC, Gibbons SM, Magis AT. 2020. Health and disease markers correlate with gut microbiome composition across thousands of people. *Nat Commun* 11:5206. <https://doi.org/10.1038/s41467-020-18871-1>.
9. Aylagas E, Borja Á, Tangherlini M, Dell'Anno A, Corinaldesi C, Michell CT, Irigoien X, Danovaro R, Rodríguez-Ezpeleta N. 2017. A bacterial community-based index to assess the ecological status of estuarine and coastal environments. *Mar Pollut Bull* 114:679–688. <https://doi.org/10.1016/j.marpolbul.2016.10.050>.
10. Ellegaard KM, Engel P. 2016. Beyond 16S rRNA community profiling: intra-species diversity in the gut microbiota. *Front Microbiol* 7:1475. <https://doi.org/10.3389/fmicb.2016.01475>.
11. Rainey PB, Travisano M. 1998. Adaptive radiation in a heterogeneous environment. *Nature* 394:69–72. <https://doi.org/10.1038/27900>.
12. Behringer MG, Choi BI, Miller SF, Doak TG, Karty JA, Guo W, Lynch M. 2018. *Escherichia coli* cultures maintain stable subpopulation structure during long-term evolution. *Proc Natl Acad Sci U S A* 115:E4642–E4650. <https://doi.org/10.1073/pnas.1708371115>.
13. Poltak SR, Cooper VS. 2011. Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. *ISME J* 5: 369–378. <https://doi.org/10.1038/ismej.2010.136>.
14. Maddamsetti R, Lenski RE, Barrick JE. 2015. Adaptation, clonal interference, and frequency-dependent interactions in a long-term evolution experiment with *Escherichia coli*. *Genetics* 200:619–631. <https://doi.org/10.1534/genetics.115.176677>.
15. Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM. 2017. The dynamics of molecular evolution over 60,000 generations. *Nature* 551:45–50. <https://doi.org/10.1038/nature24287>.
16. Blount ZD, Borland CZ, Lenski RE. 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci U S A* 105:7899–7906. <https://doi.org/10.1073/pnas.0803151105>.
17. Helling RB, Vargas CN, Adams J. 1987. Evolution of *Escherichia coli* during growth in a constant environment. *Genetics* 116:349–358. <https://doi.org/10.1093/genetics/116.3.349>.
18. Yang D-D, Alexander A, Kinnersley M, Cook E, Caudy A, Rosebrock A, Rosenzweig F. 2020. Fitness and productivity increase with ecotypic diversity among *Escherichia coli* strains that coevolved in a simple, constant environment. *Appl Environ Microbiol* 86:e00051–20. <https://doi.org/10.1128/AEM.00051-20>.
19. Schick A, Kassen R. 2018. Rapid diversification of *Pseudomonas aeruginosa* in cystic fibrosis lung-like conditions. *Proc Natl Acad Sci U S A* 115: 10714–10719. <https://doi.org/10.1073/pnas.1721270115>.
20. Jorth P, Staudinger BJ, Wu X, Hisert KB, Hayden H, Garudathri J, Harding CL, Radey MC, Rezayat A, Bautista G, Berrington WR, Goddard AF, Zheng C, Angermeyer A, Brittnacher MJ, Kitzman J, Shendure J, Fligner CL, Mittler J, Aitken ML, Manoel C, Bruce JE, Yahr TL, Singh PK. 2015. Regional isolation drives bacterial diversification within cystic fibrosis lungs. *Cell Host Microbe* 18:307–319. <https://doi.org/10.1016/j.chom.2015.07.006>.
21. Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, Polz MF. 2008. Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320:1081–1085. <https://doi.org/10.1126/science.1157890>.
22. Chase AB, Weihe C, Martiny JBH. 2021. Adaptive differentiation and rapid evolution of a soil bacterium along a climate gradient. *Proc Natl Acad Sci U S A* 118:e2101254118. <https://doi.org/10.1073/pnas.2101254118>.
23. Behringer MG, Ho W-C, Miller SF, Meraz JC, Boyer GF, Lynch M. 2020. Evolutionary response to resource deprivation: parallelism and nonmonotonicity. *bioRxiv* 865584. <https://www.biorxiv.org/content/10.1101/865584v2.full>.
24. Ismail WM, Tang H. 2019. Clonal reconstruction from time course genomic sequencing data. *BMC Genomics* 20:1002. <https://doi.org/10.1186/s12864-019-6328-3>.
25. Bawn M, Hernandez J, Trampari E, Thilliez G, Webber MA, Kingsley RA, Hall N, Macaulay IC. 2020. Bacterial single-cell genomics enables phylogenetic analysis and reveals population structures from in vitro evolutionary studies. *bioRxiv* 2020.08.25.266213. <https://www.biorxiv.org/content/10.1101/2020.08.25.266213v2.full>.
26. D'hoë K, Vet S, Faust K, Moens F, Falony G, Gonze D, Lloréns-Rico V, Gelens L, Danckaert J, De Vuyst L, Raes J. 2018. Integrated culturing, modeling and transcriptomics uncovers complex interactions and emergent behavior in a three-species synthetic gut community. *Elife* 7:e37090. <https://doi.org/10.7554/eLife.37090>.
27. Blattman SB, Jiang W, Oikonomou P, Tavazoie S. 2020. Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. *Nat Microbiol* 5: 1192–1201. <https://doi.org/10.1038/s41564-020-0729-6>.
28. Kuchina A, Brettner LM, Paleologu L, Roco CM, Rosenberg AB, Carignano A, Kibler R, Hirano M, DePaolo RW, Seelig G. 2021. Microbial single-cell RNA sequencing by split-pool barcoding. *Science* 371:eaba5257. <https://doi.org/10.1126/science.aba5257>.
29. Imdahl F, Vafadarnejad E, Homberger C, Saliba A-E, Vogel J. 2020. Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. *Nat Microbiol* 5:1202–1206. <https://doi.org/10.1038/s41564-020-0774-1>.
30. van Boxtel C, van Heerden JH, Nordholt N, Schmidt P, Bruggeman FJ. 2017. Taking chances and making mistakes: non-genetic phenotypic heterogeneity and its consequences for surviving in dynamic environments. *J R Soc Interface* 14:20170141. <https://doi.org/10.1098/rsif.2017.0141>.