

Ecological Stochasticity in Subsurface Microbial Community Assembly under Stress Gradient: Application of A General Quantitative Framework

D. Ning^{1*} (ningdaliang@ou.edu), Y. Deng,¹ J.D. Van Nostrand,¹ L. Wu,¹ P. Zhang,¹ Z. He,¹ Y. Fu,¹ D.J. Curtis,¹ Y. Li,¹ Y. Fan,¹ M.B. Smith,² A.M. Rocha,³ C.S. Smillie,² S.W. Olesen,² C.J. Paradis,⁴ J.H. Campbell,^{3,12} J.L. Fortney,⁴ T.L. Mehlhorn,³ K.A. Lowe,³ J.E. Earles,³ J. Phillips,³ S.M. Techtmann,⁴ D.C. Joyner,⁴ K.L. Bailey,³ R.A. Hurt Jr.,³ S.P. Preheim,² M.C. Sanders,² J. Yang,² M.A. Mueller,³ W.A. Lancaster,⁵ B.J. Vaccaro,⁵ F.L. Poole II,⁵ S. Brooks,³ D.B. Watson,³ A. Aaring,⁶ B. Adams,⁴ S. Brewer,⁴ K. De Leon,⁷ K. Fitzgerald,³ G.X. Ge,⁵ C. Hans,⁶ S. Kosina,⁶ L. Lui,⁶ E. Majumder,⁷ J.-W. Moon,³ A. Ottwell,⁸ S. Pfiffner,⁴ H. Smith,⁹ M. Thorgersen,⁵ S. Turkarslan,⁸ F. von Netzer,¹⁰ D. Williams,⁴ S.X. Wu,⁶ G. Zane,⁷ A. Zelaya,⁹ E.J. Alm,² N.S. Baliga,⁸ A.M. Deutschbauer,⁶ M.W. Fields,⁹ T.C. Hazen,^{3,4} T.R. Northen,⁶ J.D. Wall,⁷ M.W.W. Adams,⁵ R. Chakraborty,⁶ J.-M. Chandonia,⁶ D.A. Elias,³ D.A. Stahl,¹⁰ P.J. Walian,⁶ J. Zhou^{1,6}, **A.P. Arkin,^{6,11} and P.D. Adams,^{6,11}**

¹University of Oklahoma, Norman; ²Massachusetts Institute of Technology, Cambridge; ³Oak Ridge National Laboratory, Oak Ridge, TN; ⁴University of Tennessee, Knoxville; ⁵University of Georgia, Athens; ⁶Lawrence Berkeley National Laboratory, Berkeley, CA; ⁷University of Missouri, Columbia; ⁸Institute for Systems Biology, Seattle, WA; ⁹Montana State University, Bozeman; ¹⁰University of Washington, Seattle; ¹¹University of California at Berkeley, Berkeley; ¹²Northwest Missouri State University, Maryville; ¹³Michigan State University, East Lansing.

<http://enigma.lbl.gov/>

Project Goals: ENIGMA (Ecosystems and Networks Integrated with Genes and Molecular Assemblies) uses a systems biology approach to understand the interaction between microbial communities and the ecosystems that they inhabit. To link genetic, ecological, and environmental factors to the structure and function of microbial communities, ENIGMA integrates and develops laboratory, field, and computational methods. Our goal is to understand how human activity associated with energetic processes - in particular, contamination and climate change - is affecting the ecology of critical soil, groundwater, and aquifer systems.

Despite decades of debate, many fundamental questions regarding the ecological drivers of community assembly remain unanswered. Classical stress gradient theory in ecology discussed the role of different deterministic forces (abiotic filtering, biotic interactions, etc.) under different degrees of stresses, however, we still know little about the basic question how deterministic versus stochastic forces vary along stress gradients, particularly in microbial ecology. The groundwater in the Oak Ridge Integrated Field Research Challenge site (FRC, Oak Ridge, TN) has large geochemical gradients and has been comprehensively surveyed, providing a rare opportunity to examine ecological processes and drivers shaping subsurface microbial diversity. Groundwater samples were taken from 98 wells that covered the geochemical diversity across

the site. Collected samples and wells were analyzed for 205 environmental (spatial and geochemical) variables. Aliquots (4 L) were filtered through 10- μm and 0.2- μm pore size filters, DNA was extracted from 0.2- μm filters and the 16S rRNA genes were sequenced on an Illumina MiSeq sequencer. We have applied various approaches to further disentangle the mechanisms controlling community assembly (stochasticity vs determinism) based on this dataset. Although some interesting trend of ecological stochasticity are found using previous reported approaches, the algorithms have inherent limitations and the accuracy of the estimation was doubtful. To address the challenge, we proposed a general mathematical framework to provide quantitative assessment of ecological stochasticity under different situations in which deterministic factors drive the communities more similar or dissimilar than null expectations. We developed a new normalized index, followed by testing it with simulated communities considering abiotic filtering, biotic interactions, environmental noise, and spatial scales. Comparing to previous approaches, the new index (NST) showed obviously higher accuracy and precision, of which the coefficients were over 0.9 in most simulated scenarios. However, all approaches showed limited performance at large spatial scale or under very high environmental noise. We then applied the new index to an empirical study on groundwater microbial community succession in response to emulsified vegetable oil (EVO) injection at FRC, with expected trend of ecological stochasticity has been supported by various evidences. The new index revealed that community assembly processes were shifted from deterministic to highly stochastic post-EVO input, and that, as EVO is consumed, the groundwater communities gradually returned to be more deterministic similar to pre-EVO injection. Null model algorithms and community similarity metrics showed strong effects on quantitatively estimating ecological stochasticity, among which preferred algorithm and metrics were suggested based on reasonability of the results. Then, we applied this new index to investigate how ecological stochasticity varies along stress gradients at FRC. The results suggested obvious decrease of ecological stochasticity by the increase of environmental stress. Furthermore, we explored depth profile of ecological stochasticity based on a pilot study of sediment bacterial communities in a contaminated well and a background well. We found obvious variation of ecological stochasticity from vadose to saturated layers related to both selection and dispersal limitation. Across different empirical datasets we tested, the new index generally can correct the overestimation of stochasticity by previous approaches to some extent, and revealed the obvious effects of environmental stress on the role of ecological stochasticity in governing underground microbiome.

This material by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies (<http://enigma.lbl.gov>), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231