

ENIGMA: Assembly Mechanism of Subsurface Microbial Community under Stress Gradient and Adaptation of Super Phylum *Patescibacteria* with Genome Simplicity

D. Ning^{1*} (ningdaliang@ou.edu), R. Tian¹, J.D. Van Nostrand¹, L. Wu¹, P. Zhang¹, W. Shi¹, L. Wu¹, Y. Zhang¹, Y. Yang¹, D.J. Curtis¹, Z. He¹, M.B. Smith², A.M. Rocha³, C.S. Smillie², S.W. Olesen², C.J. Paradis⁴, J.H. Campbell^{3,9}, J.L. Fortney⁴, T.L. Mehlhorn³, K.A. Lowe³, J.E. Earles³, J. Phillips³, S.M. Techtmann⁴, D.C. Joyner⁴, D.A. Elias³, K.L. Bailey³, R.A. Hurt Jr.³, S.P. Preheim², M.C. Sanders², J. Yang², M.A. Mueller³, W.A. Lancaster⁵, B.J. Vaccaro⁵, F.L. Poole II⁵, M.W. Fields⁷, E.J. Alm², T.C. Hazen^{3,4}, M.W.W. Adams⁵, P.D. Adams^{6,8}, A.P. Arkin^{6,8}, J. Zhou^{1,6}

¹The University of Oklahoma; ²Massachusetts Institute of Technology; ³Oak Ridge National Laboratory; ⁴University of Tennessee; ⁵University of Georgia; ⁶Lawrence Berkeley National Laboratory; ⁷Montana State University; ⁸University of California at Berkeley; ⁹Northwest Missouri State University.

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

Project Goals: ENIGMA -Ecosystems and Networks Integrated with Genes and Molecular Assemblies mission is to support the development of laboratory and computational tools that link the molecular functions within individual microbes to the integrated activities of microbial communities as they interact with their environment. 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.

Community assembly mechanism under a stress gradient is a fundamental question in ecology. Classical stress gradient theory discussed the roles of different deterministic forces under different degrees of stresses, however we still know little about the relative importance of various stochastic and deterministic processes along stress gradients. 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. The 16S rDNA were sequenced for all samples by Illumina MiSeq and metagenomic shotgun sequencing were performed for 12 representative samples by Illumina HiSeq. We applied various approaches to disentangle the ecological processes controlling community assembly. Since obvious limitation of previous methods, we developed two new approaches to infer community assembly mechanisms. The first one is a general framework to estimate ecological stochasticity with a new index, normalized stochasticity ratio (NST). Another one is a quantitative framework to estimate relative influence of major ecological processes, respecting the fact that different microorganisms can be under different assembly mechanisms. We tested the index with simulated communities and demonstrated substantial improvement of quantitative

and qualitative performance of the new approaches. We applied these new approaches to investigate how ecological stochasticity and different ecological processes varied along stress gradients at FRC. The results suggested a decrease of ecological stochasticity with the increase of environmental stress, consistent with multivariate analysis. The influence of heterogeneous selection sharply increased around 3 fold from low-stress to high-stress wells, related to chemical stresses imposed by abnormal pH, nitrate, carbon limitation, uranium, and some other metals. Dispersal limitation appeared as the most influential process (42% on average), corresponding to the strong limitation of microbial migration in groundwater system. Dispersal limitation becomes less important under higher stresses, in accordance with the connectivity of groundwater among highly contaminated wells, and thus significantly correlated with pH and some metals in the supernatant rather than pellet-associated metals. In addition to community assembly mechanism variation, we also observed members from an uncultivated superphylum *Patescibacteria*, which was found prevalent in subsurface environments, but the mechanisms for its prevalence are not known. The genomic features and metabolisms of this super phylum were investigated through genome-resolved metagenomics analysis. While the members of *Patescibacteria* had reduced genomes (~1 Mbp) exclusively, our results demonstrated they retained functions essential to growth and reproduction such as genetic information processing. Surprisingly, they have sharply reduced redundant and unessential functions of metabolisms, cellular activities and stress responses. The *Patescibacteria* have ultra-small cell sizes and simplified membrane structures including flagellar assembly, transporters and two-component systems. Despite the lack of CRISPR viral defense, the bacteria could have alternative strategies to evade predation such as lacking phage receptors in membranes, which may explain the lack of phage-related proteins detected in the genomes. By establishing the linkages between bacterial features and the groundwater environmental conditions, we noticed that the adaptation of *Patescibacteria* to the environment could drive the features of a reduced genome, ultra-small cell size and lack of CRISPR viral defense. Altogether, the contamination stress gradient significantly switched the microbial community assembly mechanism and the special environment of FRC groundwater led to adaption of the superphylum *Patescibacteria* with genome simplicity.

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