

Genome-resolved metagenomics as a foundation for understanding of biogeochemistry

Jill Banfield^{1,2,*}, Karthik Anantharaman¹, Peter Andeer², Tyler Arbour¹, Christopher Brown¹, David Burstein¹, Cindy Castelle¹, Spencer Diamond¹, Sirine Fakra², Mary Firestone¹, Benjamin Gilbert², Robert Hettich³, Susan Hubbard³, Rose Kantor¹, Adi Lavy¹, Paula Matheus Carnevali¹, Trent Northen³, Chongle Pan³, Alexander Probst¹, Evan Starr¹, Brian Thomas¹, Lesley Warren⁴, Ken Williams³

¹The University of California, Berkeley, USA

²Lawrence Berkeley National Laboratory, USA

*correspondence: jbfield@berkeley.edu

³Oak Ridge National Laboratory, USA

⁴University of Toronto, Canada

Decades of research on isolated microorganisms laid the foundation for our current understanding of biogeochemical cycling. Lacking was information about the metabolic roles of the majority of organisms that were unknown or not amenable to laboratory study. We have begun to address this by integrating multi-omics measurements, geochemical data, isotopic labelling, microscopic observations and synchrotron analyses. We find that organisms with complex interdependencies arising from symbiotic lifestyles and metabolic handoffs substantially shape biogeochemical processes. Our analyses relied upon thousands of high quality genomes that also provided a new view of bacterial and archaeal diversity. Knowledge gaps relate to processes mediated by proteins of unknown function. Research to date provides the foundation for future microbial consortia studies coupled to novel genome editing and metabolic assays needed to assign functions to these novel proteins. By this approach, we have set the path to realize a more comprehensive understanding of processes that shape the geochemistry of the terrestrial subsurface.