Diversity and Metabolic Potential of the Terrestrial Subsurface Microbiome and its Influence on Biogeochemical Cycling

Jillian Banfield (jbanfield@berkeley.edu) - UC Berkeley (PI), Cindy J. Castelle, Itai Sharon, Laura A. Hug, Christopher T. Brown, Rose S. Kantor, Brian C. Thomas, Andrea Singh, Michael J. Wilkins, PNNL; Robert L. Hettich, ORNL; Susannah G. Tringe, LBNL; Kenneth H. Williams, LBNL; Eoin L. Brodie, LBNL; Harry R. Beller, LBNL; Susan S. Hubbard, LBNL (Co-Pis).

We are developing approaches for quantifying and simulating the interactions that govern biogeochemical cycling in terrestrial subsurface environments at genome-through-watershed scales (LBNL Sustainable Systems SFA 2.0). The terrestrial subsurface is a highly complex component of the Earth System and its biology is largely unknown. In fact, microbially mediated subsurface processes may be the least understood aspect of terrestrial biogeochemical cycling. Our approach to analysis of the microbial composition and metabolic potential of the subsurface leverages cultivation-independent DNA sequencing methods that are being applied to samples of sediment, groundwater, and field manipulation experiments from a floodplain site adjacent to the Colorado River, USA. Samples are collected from a range of depths, aquifer locations, and time periods. Our objective is to reconstruct draft and curated genomes from very large sequence datasets in order to predict the metabolic potential of each organism, in some cases coupled with proteomic and transcriptomic analyses and isolate-based experiments. Our studies have uncovered a remarkably high degree of taxonomic novelty within both the bacterial and archaeal domains and shown relatively little overlap in community composition across sampling sites. Notable is the finding of relatively unvarying metabolic potential within a radiation of many candidate phyla (CP; phyla for which no single member has been isolated). Based on more than 700 genomes from 11 of these lineages (including 12 complete, finished genomes) we predict small genome size, small cell size, low metabolic rates, limited metabolic potential, and a fermentation-based lifestyle. Based on many draft and two complete archaeal genomes, we have also defined a major subdivision within the archaeal domain; this sub-domain includes all known nanoarchaea. Genomes from the CP bacteria and archaea and for hundreds of organisms from many other lineages, encode a wide variety of metabolic traits, including the ability to fix carbon (in some cases, via novel RuBisCO-mediated pathways) and nitrogen, to reduce nitrate and/or nitrite, to oxidize sulfur, reduce sulfate, produce short-chain fatty acids, and to produce and/or use hydrogen. This information is being used, in the context of experiments and geochemical information, as inputs for the development of metabolic models to describe interactions and elemental cycling and its relation to hydrologic and geochemical fluctuations.