Sustainable Systems SFA 2.0: Extensive Genome-Resolved Analyses Reveal Roles for Uncultivated Archaea and Bacteria in Subsurface Biogeochemical Cycling

Jillian Banfield (jbanfield@berkeley.edu) - UC Berkeley (PI), Harry R. Beller (LBNL); Eoin L. Brodie (LBNL); Mike J. Wilkins (PNNL); Kenneth H. Williams (LBNL) (Co-PIs).

The subsurface is a vast reservoir of organic compounds and hosts a large fraction of microbial life, yet relatively little is known about sediment-associated and groundwater communities or their roles in carbon, nitrogen, sulfur and H2 cycling. The Sustainable Systems Scientific Focus Area 2.0 (SFA) is investigating subsurface carbon biogeochemical cycling, and how these processes are impacted by changes in climate and land-use. The "Metabolic Potential" component is characterizing pathways in subsurface microbial communities to inform the central modeling effort of the SFA: Genome-Enabled Watershed Simulation Capability (GEWaSC). We have conducted deep metagenome sequencing of both unamended and acetate-stimulated sediment and groundwater samples from multiple depths and locations in an alluvial aquifer adjacent to the Colorado River, CO, USA. The genomic dataset comprises over 2.5 Tb of short-read Illumina sequences and three long-read Moleculo datasets. Notably, of the thousands of genomes recovered, almost all are novel at the genus-level or higher with little overlap in community membership across sample locations or matrix types. Many of the sampled archaea and bacteria are affiliated with phyla lacking cultivated representatives and some define previously unrecognized phyla. For example, we resolved three new archaeal phyla based on sequences from 111 newly sampled genomes and reconstructed two complete (finished) genomes. For bacteria, we sampled multiple new phylum-level lineages, reconstructed hundreds of draft genomes for two specific candidate phyla, recovered around two hundred near-complete genomes from across the bacterial domain, and finished twelve genomes. Based on genome size, we infer that many of the novel archaea and bacteria have ultrasmall cell size. Detailed metabolic analyses provided evidence for possible symbiotic lifestyles. The lack of recognizable respiratory pathways involved in sulfur, nitrogen, methane, oxygen or metal cycling suggests that the biogeochemical impact of these ultrasmall organisms is primarily associated with carbon fermentation and H2 cycling. Metabolic predictions from the remaining members across all genomic datasets identify organisms possessing key traits such as nitrogen fixation, nitrate reduction, sulfur oxidation, H2 production/consumption, degradation of refractory organic compounds and carbon fixation. This information, in combination with data regarding genome GC content, codon usage and genome size, constrain metabolic rates and define functional guilds as inputs to GEWaSC. Overall, our results have dramatically expanded genomic sampling of the microbial world, improved the taxonomic resolution of candidate phyla, and augmented understanding of metabolism within domains Archaea and Bacteria.