Use of metagenomic and meta-transcriptomic analysis to interpret biogeochemical processes mediated by Hanford 100H aquifer bacteria (lab and field studies)

LBNL SFA (Laboratory Research Manager: Susan Hubbard)

H. R. Beller (PI), R. Han, U. Karaoz, H. C. Lim, L. Yang, B. Faybishenko, E. L. Brodie—*LBNL*

We are developing the use of biomolecular signatures as part of the LBNL SFA challenge entitled "Unraveling Biogeochemical Reaction Networks Mediating Sustained Chromium Reduction", which focuses on *in situ* reductive immobilization of Cr at DOE's Hanford 100H site. By applying a high-throughput approach that uses metagenomic and meta-transcriptomic data, we hope to identify highly expressed genes in a specific microbial community under conditions of interest, without requiring any *a priori* sequence information or assumptions about what processes might be occurring. A key scientific goal is to determine if there are diagnostic biomolecular signatures indicative of important aquifer biogeochemical processes that can be used to (a) help discriminate between direct (enzymatic) and indirect (abiotic) oxidation-reduction processes relevant to bioremediation and (b) to inform and constrain reactive transport models even when geochemical field measurements do not reveal all relevant processes. We have collected metagenome and meta-transcriptome sequence information from various experimental systems (including flow-through columns and batch microcosms inoculated with Hanford aquifer sediment or groundwater, bacterial isolates from the flow-through columns, and site groundwater) under conditions relevant to *in situ* chromate reduction and re-oxidation at Hanford 100H. Overall, 12 metagenome and 14 meta-transcriptome samples were sequenced, ranging from 0.1 to 1.6 Gb of sequence data per sample. This poster gives an overview of several noteworthy findings from these Hanford studies.

In microcosm studies designed to characterize functional changes in an aquifer-derived, chromate-reducing microbial community as it transitions successively through electron-accepting conditions relevant to the Hanford subsurface, one condition tested was nitrate-dependent oxidation of Fe(II) and reduced S compounds. Meta-transcriptome sequence data revealed expression of nitrate reductase (*nar*) and S oxidation (*sox*) gene clusters from an alpha-Proteobacterial species, whereas routine geochemical (ion chromatography) data did not yield evidence of the S oxidation product. Thus, the meta-transcriptomic data provided insights into biogeochemical processes that would not have been detected by routine geochemical analysis.

In flow-through columns containing Hanford aquifer sediment and reducing Cr(VI) under denitrifying conditions, meta-transcriptome data revealed complete denitrification (with expression of *nar*, *nir*, *nor*, and *nos* genes, as well as the highly expressed pseudoazurin), but the most highly expressed genes encoded proteins of unknown function. This finding stresses an advantage of metagenomic and meta-transcriptomic sequencing over *a priori* approaches, namely, that sequencing can reveal important but unannotated genes/enzymes and focus research on such knowledge gaps. In addition, we isolated a Firmicute (strain HCF1) that dominated the microbial communities of flow-through columns that rapidly reduced Cr(VI) under fermentative conditions; we were able to metabolically reconstruct key pathways (such as lactate fermentation) by examining the genome sequence of strain HCF1.

Metagenome and meta-transcriptome sequencing was carried out on groundwater samples collected during various biogeochemical regimes (e.g., denitrification, sulfate reduction) that followed lactate biostimulation at the 100H site. Among other findings was evidence of nitrogen and phosphorous limitation, which provided information on amendments needed to optimize *in situ* biostimulation.

Following this biomolecular signature discovery phase that involves metagenome and meta-transcriptome sequencing of a range of Hanford-relevant biogeochemical conditions, we are developing meta-transcriptome-based gene expression microarrays that we hope will allow us to rapidly profile the *in situ* gene expression of Hanford microbial communities on a qualitative basis (i.e., what pathways are most highly expressed?) and a semi-quantitative basis.