ENIGMA Biotechnology: Metabolic Profiling of Bacterial Mutant Libraries to Link Metabolite Utilization to Genotype

Richard Baran* (RBaran@lbl.gov), Benjamin Bowen, Morgan Price, Adam A. Arkin, Adam Deutschbauer, **Trent Northen**

Lawrence Berkeley National Lab, Berkeley CA

http://enigma.lbl.gov/

Project goals: Microorganism-based approaches to address DOE mission goals in remediation, carbon sequestration and energy production will require quantitative understanding of biological complexity at multiple scales – from molecular networks of individual species to the dynamic inter-species interactions within the communities in which they reside. The broad goals of ENIGMA are to understand, at a molecular systems level, the bacterial soil communities at DOE sites contaminated with heavy metals or radionuclides with sufficient detail to predictively model interactions within microbial and community processes that drive complex geochemistry in key environments. In doing so we expect to define biological principles governing selection of microbial community function and composition in given environments.

The discrepancy between the pace of sequencing and functional characterization of genomes has been recognized as one of the major challenges in microbial genomics. The ENIGMA program at LBNL is addressing this by integrating high-throughput genetics and mutant libraries with metabolic observations to establish direct functional associations between genes and metabolic processes. Mass spectrometry (MS) based metabolomics allows the profiling of metabolites in complex biological samples with high sensitivity and is well suited for interrogating the metabolic capabilities of microorganisms. We have used untargeted metabolomics to identify unexpected and novel metabolites as well as to map the uptake and release of a broad range of metabolites by different microorganisms. Large, genome-wide libraries of mutant strains, developed for multiple ENIGMA microorganisms, enable investigations into the role of specific genes in various metabolic processes using high-throughput metabolomics. However, metabolite profiling of complex samples is usually performed using time-consuming chromatography to separate the metabolites prior MS analysis. This makes it largely incompatible with screening large mutant libraries. To meet this challenge, we developed a workflow combining untargeted and high-throughput metabolomics to identify genes of *Escherichia coli* and *Shewanella oneidensis* MR-1 related to the utilization of specific metabolites. The ability to uptake specific metabolites from different complex media was analyzed using liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). A subset of metabolites found to be utilized were supplemented to minimal media. A limited complexity of the media facilitated high-throughput analysis of spent media extracts using MS. A total of 8000 mutant strains were screened for defects in metabolite utilization. The presence of one of the tested metabolites in the spent media extracts of specific mutants directly links genes to a metabolic defect. Intracellular metabolites of these mutants were then profiled by LC-MS to identify potential accumulation of intermediates related to the utilization of specific metabolites. Using this approach, we identified genes of known function as well as putative transport proteins and enzymes with previously ambiguous annotations.

This work conducted by ENIGMA was supported by the Office of Science, Office of Biological and Environmental Research, of the U. S. Department of Energy under Contract No. DE-AC02-05CH11231.