Evidence-based annotation of gene function in metal-reducing and sulfate-reducing bacteria using genome-wide mutant fitness data

Adam Deutschbauer[1] (AMDeutschbauer@lbl.gov), Grant M. Zane[2], Samuel Fels[2], Morgan Price[1], Jennifer V. Kuehl[1], Kelly Wetmore[1], Swapnil R. Chhabra[1], Thomas R. Juba[2], Gareth P. Butland[1], **Adam P. Arkin**[1], **Judy D. Wall**[2] and Paul Adams[1]

[1]Lawrence Berkeley National Lab, Berkeley, CA, [2]Univ. of Missouri, Columbia, MO

http://enigma.lbl.gov

Project Goals: The ENIGMA (Ecosystems and Networks Integrated with Genes and Molecular Assemblies) project seeks to elucidate the mechanisms and key processes that enable microorganisms and their communities to function in metal-contaminated soil sites. One goal is to develop and apply high-throughput genetic technologies to identify causative genes and pathways rapidly in ENIGMA-relevant microorganisms.

Abstract: The development of genetic tools to study the functions of thousands of poorly characterized genes in newly isolated environmental microbes is a rate-limiting step in determining the activity of the keystone organisms in communities. One promising approach to meet this challenge is the large-scale generation of mutant libraries and gene fitness assays as a complement to standard whole genome sequencing and classical annotations based on homology and computation. Currently, large-scale mutant collections have been assembled for only a small sample of typically pathogenic bacteria. With the availability of a universal, sequence-verified TagModule collection (Oh et al. NAR 38:e146, 2010) that has been introduced into a hyperactive mini-Tn*5* (Larsen et al., Arch Microbiol. 178:193, 2002), the creation of such libraries has been facilitated for a number of environmental microbes. We have generated sequenced and archived transposon mutant collections of 24,688 mutants of *Shewanella oneidensis* MR-1, 15,477 of *Desulfovibrio alaskensis* G20, and 13,007 of *Desulfovibrio vulgaris* Hildenborough. Simple scanning of the genes not mutated provides a limited list that includes essential genes for the condition of mutant isolation. Pools of tagged or barcoded transposon mutants of most non-essential genes can be used to probe the gene fitness in different growth conditions and when subjected to various stresses. Recent work with *S. oneidensis* (Deutschbauer et al., PLOS Genetics, 7:e1002385, 2011) has shown that many hypothetical genes and potentially redundant genes have detectable phenotypes when this approach is applied across a large number of diverse growth conditions. In some instances, specific molecular functions can be predicted (evidence-based annotations) where nearly meaningless annotations were the extent of previous knowledge.

This work conducted by ENIGMA (Ecosystems and Networks Integrated with Genes and Molecular Assemblies) 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.