Identification of Novel Biosynthetic and Catabolic Pathways in Diverse Bacteria Using High-throughput Genetics

Morgan N. Price[1], Kelly M. Wetmore[1], R. Jordan Waters[2], Mark Callaghan[1], Jayashree Ray[1], Jennifer V. Kuehl[1], Ryan A. Melnyk[1], Jacob S. Lamson[1], Yumi Suh[1], Zuelma Esquivel[1], Harini Sadeeshkumar[1], Romy Chakraborty[3], Benjamin E. Rubin[4], Grant M. Zane[5], Judy D. Wall[5], James Bristow[2], Matthew J. Blow[2], Adam P. Arkin[1,6], **Adam M. Deutschbauer[1*] (AMDeutschbauer@lbl.gov)**, Paul D. Adams[6,7]

[1]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA; [2]Joint Genome Institute, Walnut Creek, CA; [3] Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory; [4]Division of Biological Sciences, University of California, San Diego; [5]Biochemistry Division, University of Missouri, Columbia, Missouri; [6]Department of Bioengineering, University of California, Berkeley; [7]Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA

http://fit.genomics.lbl.gov/; http://enigma.lbl.gov/;

Project Goals: For most bacteria with sequenced genomes, we do not understand how they synthesize some amino acids or how they consume some carbon sources. This makes it challenging to reconstruct their metabolism or to predict their ecological role from their genome sequence. We used genome-wide mutant fitness data from 14 different genera of heterotrophic bacteria to fill gaps in amino acid biosynthesis pathways and to identify novel catabolic routes for several sugars.

The genomes of many free-living bacteria do not appear to encode some of the necessary enzymes for amino acid biosynthesis. This has led to speculation that these bacteria cannot make all 20 amino acids and that they are cross-feeding each other amino acids. However, when we tested 25 heterotrophic Proteobacteria from 14 genera, we found that 24 of them grew in minimal media without any added amino acids. In contrast, comparative genomics tools predict that all of these bacteria, except for *Escherichia coli*, are auxotrophic for multiple amino acids. We examined representatives of 10 genera in more detail and identified 11 gaps that we could not fill using current knowledge. Using genome-wide mutant fitness data, we identified novel enzymes that fill 9 of the 11 gaps and hence explain the biosynthesis of methionine, threonine, serine, or histidine by bacteria from six genera. We also found that the sulfate-reducing bacterium *Desulfovibrio vulgaris* synthesizes homocysteine (which is a precursor to methionine) by using DUF39, NIL/ferredoxin, and COG2122 proteins, and that homoserine is not an intermediate in this pathway. It appears that most free-living bacteria can make all 20 amino acids but we do not yet know how.

We also used the genome-wide fitness data to identify novel catabolic pathways for several sugars:

- Oxidation of 2-deoxy-D-ribose, a component of DNA, by *Pseudomonas simiae* (4 novel reactions)
- Oxidation of D-arabinose by *Sinorhizibium meliloti* (2 novel reactions)
- Additional dehydratase steps in the oxidation of L-fucose by *Sinorhizobium meliloti*

- The first identification of a gene for D-glucosaminate ammonia-lyase, which is required for glucosamine catabolism in many *Pseudomonas*

Funding statement. This material by ENIGMA - Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory, is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231.