

## **Strain isolation, genome sequencing, and functional genomics reveals adaptive evolution of a *Pseudomonas* population at a human impacted field site**

E.J. Alm<sup>1\*</sup>(ejalm@mit.edu), S.J. Spencer<sup>1</sup>, V.K. Mutalik<sup>2,4</sup>, T. C. Hazen<sup>5,6</sup>, J.-M. Chandonia<sup>2</sup>, J. Zhou<sup>3</sup>, R. Chakraborty<sup>2</sup>, Deutchbauer A.<sup>2</sup>, **A.P. Arkin<sup>2,4</sup>, P.D. Adams<sup>2,4</sup>**

<sup>1</sup>Massachusetts Institute of Technology; <sup>2</sup>Lawrence Berkeley National Lab; <sup>3</sup>University of Oklahoma; <sup>4</sup>University of California at Berkeley; <sup>5</sup>University of Tennessee; <sup>6</sup>Oak Ridge National Lab

<http://enigma.lbl.gov>

**Project Goals: Whole genome sequencing of environmental strains can reveal recent adaptive changes, making it a powerful tool to understand how environmental conditions impact the microbial genomes in those environments. Here we aimed to combine whole genome sequencing with functional genomics of field isolates to build a mechanistic understanding of natural selection at a human impacted groundwater site.**

Microbial communities in groundwater ecosystems are inherently difficult to study due to sampling challenges paired with highly dynamic microenvironments. A number of recent efforts have characterized and mined the vast bacterial diversity within groundwater via amplicon surveys and shotgun metagenomics, but these data sources cannot reveal dynamic or historical changes without high-resolution sampling. Here we identify recent microbial adaptation within groundwater sites using isolation and whole genome sequencing, and combined these insights with functional genomic experiments directly measuring the contribution of genes to fitness under controlled laboratory conditions.

We isolated 139 *Pseudomonas* strains, that grouped into 15 lineages, and found that each lineage contained strains from multiple sampling sites across a broad geographic region. SNP analysis confirmed that recent mutational changes were impacting strains in individual sampling sites, despite a fast-flowing aquifer, and a two-component system showed evidence of positive selection. We also searched for gene gain and loss within strain isolates, and found evidence of transcriptional gene excision within otherwise clonal isolates at a single location.

We next sought to investigate the genes identified from genome sequencing with high-throughput *in vitro* fitness measurements. Transposon-based mutagenesis was used to generate a saturating library of knock out mutants which were grown across hundreds of conditions. Fitness measurement revealed that genes in the signaling pathway identified by genome analysis as undergoing adaptation were also among the most likely to be selected for across a variety of *in vitro* conditions. These data suggest it may be possible to not only identify the genes under selection, but the environmental factors driving that selection by combining sequencing with high-throughput functional genomics.

*Funding statement: ENIGMA (<http://enigma.lbl.gov>) at LBNL supported by Office of Biological and Environmental Research US Dept of Energy Contract No: DE-AC02-05CH1123*