

Dynamics of Sequence-Discreet Bacterial Populations Inferred Using Metagenomics

Sarah Stevens¹, Matthew Bendall², Dongwan Kang², Jeff Froula², Rob Egan², Leong-Keat Chan², Susannah Tringe², Katherine McMahon², Rex Malmstrom²

¹University of Wisconsin – Madison, Department of Bacteriology

²Department of Energy Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA, USA, USA

March 2014

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

Population Dynamics of Sequence-Discrete Bacterial Populations Inferred Using Metagenomes

Sarah Stevens*¹, Matthew Bendall², Dongwan Kang², Jeff Froula², Leong-Keat Chan², Susannah Tringe², Rex Malmstrom², and Katherine McMahon¹

[*sstevens2@wisc.edu](mailto:sstevens2@wisc.edu)

Affiliations: ¹University of Wisconsin - Madison, Dept. of Bacteriology, Madison, WI, USA; ²Department of Energy Joint Genome Institute, Walnut Creek, CA, USA;

From a multi-year metagenomic time series of two dissimilar Wisconsin lakes we have assembled dozens of genomes using a novel approach that bins contigs into distinct genome based on sequence composition, e.g. kmer frequencies, and contig coverage patterns at various times points. Next, we investigated how these genomes, which represent sequence-discrete bacterial populations, evolved over time and used the time series to discover the population dynamics. For example, we explored changes in single nucleotide polymorphism (SNP) frequencies as well as patterns of gene gain and loss in multiple populations. Interestingly, SNP diversity was purged at nearly every genome position in some populations during the course of this study, suggesting these populations may have experienced genome-wide selective sweeps. This represents the first direct, time-resolved observations of periodic selection in natural populations, a key process predicted by the 'ecotype model' of bacterial diversification.