Lawrence Berkeley National Laboratory
Recent Work

Title
Genome-wide Selective Sweeps in Natural Bacterial Populations Revealed by Time-series Metagenomics

Permalink
https://escholarship.org/uc/item/4t50395d

Authors
Chan, Leong-Keat
Bendall, Matthew L.
Malfatti, Stephanie
et al.

Publication Date
2014-05-19
Genome-wide selective sweeps in natural bacterial populations revealed by time-series metagenomics

Leong-Keat Chan\textsuperscript{1}, Matthew L. Bendall\textsuperscript{1}, Stephanie Malfatti\textsuperscript{1}, Patrick Schwientek\textsuperscript{1}, Julien Tremblay\textsuperscript{1}, Wendy Schackwitz\textsuperscript{1}, Joel Martin\textsuperscript{1}, Amrita Pati\textsuperscript{1}, Brian Bushnell\textsuperscript{1}, Brian Foster\textsuperscript{1}, Dongwan Kang\textsuperscript{1}, Suannah G. Tringe\textsuperscript{1}, Stefan Bertilsson\textsuperscript{2}, Mary Ann Moran\textsuperscript{3}, Ashley Shade\textsuperscript{4}, Ryan J. Newton\textsuperscript{5}, Sarah Stevens\textsuperscript{6}, Katherine D. McMahon\textsuperscript{6}, and Rex Malmstrom\textsuperscript{1}

\textsuperscript{1} DOE Joint Genome Institute
\textsuperscript{2}Uppsala University
\textsuperscript{3}University of Georgia
\textsuperscript{4}Yale University
\textsuperscript{5}University of Wisconsin-Milwaukee
\textsuperscript{6}University of Wisconsin - Madison

May 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.
Genome-wide selective sweeps in natural bacterial populations revealed by time-series metagenomics

Leong-Keat Chan1, Matthew L. Bendall1, Stephanie Malfatti1, Patrick Schwientek2, Julien Tremblay1, Wendy Schackwitz1, Joel Martin1, Amrita Pati1, Brian Bushnell1, Brian Foster1, Dongwan Kang1, Susannah G. Tringe1, Stefan Bertilsson1, Mary Ann Moran3, Ashley Shade4, Ryan J. Newton3, Sarah Stevens3, Katherine D. McMahon4, and Rex R. Malmstrom5

1DOE Joint Genome Institute 2Uppsala University 3University of Georgia 4Yale University 5University of Wisconsin-Madison 6University of Wisconsin-Madison

ABSTRACT
Multiple evolutionary models have been proposed to explain the formation of genetically and ecologically distinct bacterial groups. Time-series metagenomics enables direct observation of evolutionary processes in natural populations, and if applied over a sufficiently long time frame, this approach could capture events such as genome-specific or genome-wide selective sweeps. Direct observations of either process could help resolve how distinct groups form in natural microbial assemblages. Here, from a three-year metagenomic study of a freshwater lake, we explore changes in single-nucleotide polymorphism (SNP) frequencies and patterns of gene gain and loss in populations of Chlorobiaceae and Methylphilaceae. SNP analyses revealed substantial genetic heterogeneity within these populations, although the degree of heterogeneity varied considerably among closely related, co-occurring Chlorobiaceae populations. SNP allele frequencies, as well as the relative abundance of certain genes, changed dramatically over time in each population. Interestingly, SNP diversity was purged at nearly every genome position in one of the Chlorobiaceae populations over the course of three years, while at the same time multiple genes either swept through or were swept from this population. These patterns were consistent with a genome-wide selective sweep, a process predicted by the ‘ecotype model’ of diversification, but not previously observed in natural populations.

APPROACH
• Shotgun sequenced freshwater community at 45 time points from 2007-2009
• Assembled 2 genomes from Chlorobiaceae and 2 from Methylphilaceae
• Mapped metagenomic reads to genomes at >95% nucleotide identity to identify: 1) ‘sequence-discrete’ populations 2) Allele frequencies at SNP loci 3) Relative gene abundance within populations

CONCLUSIONS
• The dramatic loss of SNP diversity and the patterns of gene gain and loss in the Chlorobiaceae-1280 population were consistent with a genome-wide selective sweep.
• ‘Sequence-discrete’ populations may have experienced a ‘soft sweep,’ whereas most SNP diversity in Methylotenera-330 population was lost prior to the start of this study.
• ‘Sequence-discrete’ populations behave like theoretically defined ‘ecotypes’
  - Displacement of many co-existing strains by a single strain/lineage within the same population implies that all population members shared the same ecological niche.
  - Closely related, co-occurring sequence-discrete populations experience sweeps independently

This work was supported by U.S. Dept. of Energy’s Office of Science (DE-AC02-05CH11231), and the National Science Foundation’s Microbial Observatory, Long Term Ecological Research, INSPIRE, and CAREER programs.