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

Leong-Keat Chan¹, Matthew L. Bendall¹, Stephanie Malfatti¹, Patrick Schwientek¹, Julien Tremblay¹, Wendy Schackwitz¹, Joel Martin¹, Amrita Pati¹, Brian Bushnell¹, Brian Foster¹, Dongwan Kang¹, Suannah G. Tringe¹, Stefan Bertilsson², Mary Ann Moran³, Ashley Shade⁴, Ryan J. Newton⁵, Sarah Stevens⁶, Katherine D. McMahon⁶, and Rex Malmstrom¹

¹ DOE Joint Genome Institute
² Uppsala University
³ University of Georgia
⁴ Yale University
⁵ University of Wisconsin-Milwaukee
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.
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 gene-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 Chlorobacum and Methylphilicaceae. SNP analyses revealed substantial genetic heterogeneity within these populations, although the degree of heterogeneity varied considerably among closely related, co-occurring Methylphilicaceae 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 Chlorobacum 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 times points from 2007-2009
• Assembled 2 genomes from Chlorobacum and 2 from Methylphilicaceae
• 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 Chlorobium-1280 population were consistent with a genome-wide selective sweep.
• Methylotenera-124 population 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