Analysis of an Abundant Bacterial Genus in a Leaf Litter Community

Alexander B. Chase (abchase@uci.edu), Adam C. Martiny, Ulas Karaoz, Eoin L. Brodie, and Jennifer B.H. Martiny

1University of California, Irvine, CA; 2Lawrence Berkeley National Laboratory, Berkeley, CA


Project Goals: With increasing availability of cultured and uncultured (metagenome-reconstructed and single-cell) genomes, it is now feasible to study the ecology and genetic diversity of specific microbial populations in complex communities. Previously, using 16S rRNA sequencing, we identified Curtobacterium (phylum: Actinobacteria, family: Microbacteriaceae) to be the most abundant taxon (~18% relative abundance) in a leaf litter grassland community in Southern California (Matulich et al. 2015). Here we analyze shotgun metagenomic libraries (Berlemont et al. 2014) at this site and ask: 1) How the abundance of Curtobacterium, and more specifically clades within Curtobacterium, responded over a two-year period (2010-2012) and across three treatments: drought, increased nitrogen, and control. 2) What is the genetic variation within populations of Curtobacterium over time?

We found that traditional characterization of microbial abundance in metagenomic libraries (i.e. MG-RAST) lack the genomic representation in their database to accurately quantify relevant taxa of interest, particularly within soil communities. To address these concerns, we isolated and sequenced 14 Curtobacterium genomes (Chase et al. 2016) and constructed a robust reference database and phylogenetic analyses into the Microbacteriaceae family. To quantify microbial abundance, we established an approach to extract and analyze single-copy marker genes within the metagenomic data. Normalizing the metagenomic data using 30 phylogenetically conserved marker genes, we estimated that the relative abundance of all Microbacteriaceae ranges from 22.76% to 29.12% across the samples. By placing short metagenomic reads onto this family reference tree, we showed significant strain-specific responses within clades of Curtobacterium over time. To validate these results, we then recruited and mapped reads to 4 genomes representing multiple clades. The mapping corroborated the abundance of specific clades within Curtobacterium, suggesting clade-specific responses to changing environmental conditions. Mapping of the reads further revealed genomic variation across clades, specifically with regards to the presence and/or absence of particular glycoside hydrolase genes. Together these results show that targeted analysis at fine taxonomic resolution can reveal population-specific characterizations that would otherwise be masked by broad taxonomic designations.

References


This research was funded by the Genomic Science Program, the (now retired) Program in Ecosystem Research (PER), and direct support to Lawrence Berkeley National Laboratory. The Genomic Science Program is part of the Biological System Science Division (BSSD) and PER was part of the Climate and Environmental Sciences Division (CESD) of Biological and Environmental Research (BER) in the US Department of Energy Office of Science.