

Microbial Dark Matter Analysis Using 16S rRNA Gene Metagenomics Sequences

Authors : Hana Barak, Alex Sivan, Ariel Kushmaro

Abstract : Microorganisms are the most diverse and abundant life forms on Earth and account for a large portion of the Earth's biomass and biodiversity. To date though, our knowledge regarding microbial life is lacking, as it is based mainly on information from cultivated organisms. Indeed, microbiologists have borrowed from astrophysics and termed the 'uncultured microbial majority' as 'microbial dark matter'. The realization of how diverse and unexplored microorganisms are, actually stems from recent advances in molecular biology, and in particular from novel methods for sequencing microbial small subunit ribosomal RNA genes directly from environmental samples termed next-generation sequencing (NGS). This has led us to use NGS that generates several gigabases of sequencing data in a single experimental run, to identify and classify environmental samples of microorganisms. In metagenomics sequencing analysis (both 16S and shotgun), sequences are compared to reference databases that contain only small part of the existing microorganisms and therefore their taxonomy assignment may reveal groups of unknown microorganisms or origins. These unknowns, or the 'microbial sequences dark matter', are usually ignored in spite of their great importance. The goal of this work was to develop an improved bioinformatics method that enables more complete analyses of the microbial communities in numerous environments. Therefore, NGS was used to identify previously unknown microorganisms from three different environments (industrials wastewater, Negev Desert's rocks and water wells at the Arava valley). 16S rRNA gene metagenome analysis of the microorganisms from those three environments produce about ~4 million reads for 75 samples. Between 0.1-12% of the sequences in each sample were tagged as 'Unassigned'. Employing relatively simple methodology for resequencing of original gDNA samples through Sanger or MiSeq Illumina with specific primers, this study demonstrates that the mysterious 'Unassigned' group apparently contains sequences of candidate phyla. Those unknown sequences can be located on a phylogenetic tree and thus provide a better understanding of the 'sequences dark matter' and its role in the research of microbial communities and diversity. Studying this 'dark matter' will extend the existing databases and could reveal the hidden potential of the 'microbial dark matter'.

Keywords : bacteria, bioinformatics, dark matter, Next Generation Sequencing, unknown

Conference Title : ICM 2018 : International Conference on Microbiome

Conference Location : Amsterdam, Netherlands

Conference Dates : December 03-04, 2018