Estimate of the sequenced proportion of the global prokaryotic genome

Zheng Zhang
Jianing Wang
Jinlan Wang
Jingjing Wang
Yuezhong Li

Video Byte

Keywords: Microbiome, Genome sequencing, Prokaryotic biome, Earth Microbiome Project, Predominant taxa, microbes, prokaryotes, genome

Posted Date: November 12th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-106686/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Prokaryotes are assumed to be the oldest existing life form on earth and are found in almost all ecosystems. Genome sequencing can provide a blueprint for the evolutionary and functional diversity of prokaryotes, but what proportion of bacteria and archaea on earth have actually had their genomes sequenced? A new study aimed to explore this basic question. By aligning the sequences released by the Earth Microbiome Project with 155,810 prokaryotic genomes from public databases, researchers found that the median proportions of the genome-sequenced cells and taxa in different biomes reached 38.1% and 18.8%. Due to a set of cosmopolitan OTUs that are found in multiple samples and preferentially sequenced, only 2.1% of the global prokaryotic taxa are represented by sequenced genomes, and most of the biomes were occupied by a few predominant taxa with a high relative abundance, while rarer taxa were less represented. These results demonstrate that much more remains to be learned about prokaryotes on earth and accelerate the comprehensive understanding of microbial ecological functions in different environments.