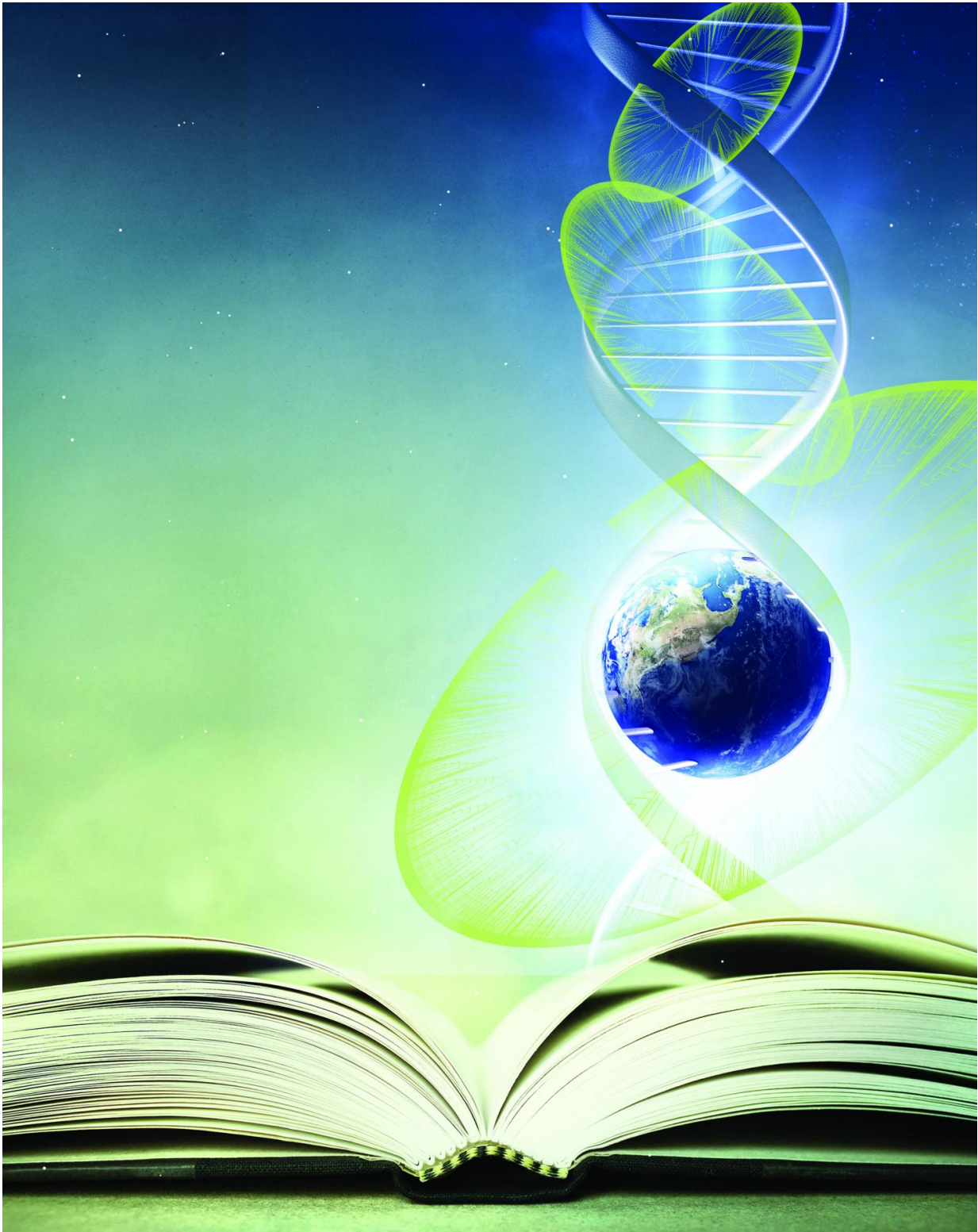


# Uncovered: 1,000 new microbial genomes

June 12 2017

---



Credit: Zosia Rostomian, Berkeley Lab Creative Services

The number of microbes in a handful of soil exceeds the number of stars in the Milky Way galaxy, but researchers know less about what's on Earth because they have only recently had the tools to deeply explore what is just underfoot. Now scientists at the U.S. Department of Energy Joint Genome Institute (DOE JGI), a DOE Office of Science User Facility, have taken a decisive step forward in uncovering the planet's microbial diversity. In a paper published online June 12, 2017 in *Nature Biotechnology*, DOE JGI's Prokaryotic Super Program head Nikos Kyrpides and his team of researchers report the release of 1,003 phylogenetically diverse bacterial and archaeal reference genomes—the single largest release to date.

"Bacteria and archaea comprise the largest amount of biodiversity of free-living organisms on Earth," said Kyrpides, senior author of the paper. "They have already conquered every environment on the planet, so they have found ways to survive under the harshest of conditions with different enzymes and with different biochemistry."

The U.S. Department of Energy is interested in learning more about this biodiversity because microbes play important roles in regulating Earth's biogeochemical cycles—processes that govern nutrient circulation in terrestrial and marine environments, for example. Uncovering the functions of genes, enzymes and metabolic pathways through [genome sequencing](#) and analysis has wide applications in the fields of bioenergy, biomedicine, agriculture and environmental sciences.

## **New Functions, New Applications**

The effort is part of the DOE JGI's Genomic Encyclopedia of Bacteria and Archaea (GEBA) initiative that aims to sequence thousands of bacterial and archaeal genomes to fill in unexplored branches of the tree of life. "In addition to identifying over half a million new protein families, this effort has more than doubled the coverage of phylogenetic

diversity of all type strains with [genome](#) sequences", said Supratim Mukherjee, a DOE JGI computational biologist and co-first author of the paper.

Since a great portion of research in microbial genomics has been focused on human pathogens or biotechnological work horses, GEBA is the main effort worldwide attempting to address the phylogenetic coverage knowledge gap by sequencing a diverse set of cultured but poorly characterized microbial type strains. "It was recognized that we weren't sampling many parts of the tree of life," said Rekha Seshadri, a DOE JGI computational biologist and co-first author of the paper. "And if we sampled some of those parts of the tree, we'd discover new functions, which could be an important resource for new applications."

The release of these genomes is the culmination of almost a decade's worth of work, with the first 56 GEBA genomes published in 2009. The microorganisms were isolated from environments ranging from sea water and soil, to plants, and to cow rumen and termite guts. Genome sequencing and analysis was done at the DOE JGI through the Community Science Program, and the 1,003 genomes are publicly available through the Integrated Microbial Genomes with Microbiomes (IMG/M) system, with all associated metadata in compliance with the Genomics Standards Consortium available through the Genomes OnLine Database. In fact, all these genomes were publicly released immediately after sequencing to maximize their use by the larger scientific community, in accordance with the DOE JGI's practice of immediate data release said co-author Tanja Woyke, head of the DOE JGI Microbial Genomics Program, who overviewed the sequencing of the project.

With the release of high quality genomic information from the 1,003 reference genomes, DOE JGI is providing a wealth of new sequences that will be invaluable to scientists interested in experiments such as

characterizing biotechnologically relevant secondary metabolites or studying enzymes that work under specific conditions, Seshadri said. And because Kyrpides' research team sequenced type strains that are readily available from culture collections, scientists can perform follow-up experiments with them in the lab, she adds. "The partnership with culture collection centers such as the Leibniz Institute DSMZ in Germany and the ATCC Global Bioresource Center in the U.S., was critical in accomplishing this endeavor," said Kyrpides.

Though it's evident that bacteria can jumpstart innovations in biotechnology—such as the species *Streptococcus pyogenes*, which produces the Cas9 protein that functions as the "scissors" in the breakthrough CRISPR-Cas9 gene editing tool—scientists have only just begun to uncover the hidden potential that exists within the wide genetic diversity of bacterial and archaeal phyla.

## **A Reference Framework to Anchor Data**

Jonathan Eisen, a microbiologist at the University of California, Davis who initiated the GEBA project at the DOE JGI in 2007 with Kyrpides and Phil Hugenholtz, and Hans-Peter Klenk at the Leibniz Institute DSMZ, believes that the paper reinforces that having a goal to achieve phylogenetic diversity is a more useful approach than random selection when choosing microbial organisms for sequencing.

He said filling out the tree of life will provide researchers with a reference framework with which to understand their own results. "It's incredibly helpful for interpreting environmental data. For example, if you go and find a fossil bed somewhere and find tons of bones, but if no one had ever assembled skeletons before, it'd be useless," Eisen said. "But with an assembled skeleton to use as a reference, 'you can say 'this looks like a mammal'. The same is true with metagenomic data—if you have reference genomes from across the tree [of life], you can anchor

environmental data much more accurately."

"At a time when we are witnessing the public databases being flooded by an infusion of low or questionable quality, highly fragmented and chimeric or contaminated genomes, the significance of genomes from the type strains as invaluable taxonomic signposts cannot be overstated," Kyrpides said.

**More information:** "1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life," *Nature Biotechnology* (2017). [DOI: 10.1038/nbt.3886](https://doi.org/10.1038/nbt.3886)

Provided by DOE/Joint Genome Institute

Citation: Uncovered: 1,000 new microbial genomes (2017, June 12) retrieved 18 April 2024 from <https://phys.org/news/2017-06-uncovered-microbial-genomes.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.