Title
Discovering untapped biochemistry in microorganisms from hidden corners.

Journal
Nature reviews. Microbiology, 20(11)

ISSN
1740-1526

Authors
Zhao, Zhiying
Yoshikuni, Yasuo

Publication Date
2022-11-01

DOI
10.1038/s41579-022-00795-w

Peer reviewed

NEWS & ANALYSIS

GENOME WATCH

Discovering untapped biochemistry in microorganisms from hidden corners

Zhiying Zhao and Yasuo Yoshikuni

This Genome Watch explores how metagenomics and bioinformatics can aid to uncover the prevalence and biosynthetic capabilities of underexplored microorganisms in various ecosystems.

Microbial genes encode much of the functional repertoire of Earth's life, and discovering their encoded biosynthetic capabilities is crucial to synthetic biology for the creation of a sustainable bioeconomy. However, most microorganisms outside laboratories cannot be cultivated, so exploiting their biochemical resources is challenging. Using artificial intelligence, Coelho et al.[1] constructed nearly 100,000 bacterial genomes and a non-redundant gene catalogue of 303 million species from 84,029 high-quality genomes and 13,174 metagenomes in 14 habitats in order to help researchers determine global microbial gene distribution and understand related protein functions. Most of the catalogued genes were habitat-specific, which suggests that their unique genomes enable the bacteria to live in distinct geographical areas and could be abundant sources of new genes and biosynthetic gene clusters (BGCs).

Paoli et al.[2] studied the global ocean microbiome. From 1,038 publicly available marine metagenomes from 215 sites worldwide, and another 10,000 genomes already documented, they reconstructed the genomes of ~26,000 marine microorganisms and identified ~2,700 undescribed species. The database enables large-scale mining of ~40,000 putative BGCs potentially producing ~7,000 mostly new, natural products. Several of these putative BGCs were in phylogenetic groups not previously thought to carry BGCs. For example, the authors unveiled a new family (Candidatus Eudoremicrobiaceae) that belongs to an uncultivated bacterial phylum, which had been thought to exist only in terrestrial environments. Characterization of two predicted BGCs indicated that a wealth of natural products and pathways may be investigated productively.

To address the fact that information from glacial habitats is limited, Liu et al. sequenced and assembled 85 metagenomes and obtained 2,358 metagenome-assembled genomes from 21 glaciers on the Qinghai–Tibet Plateau[3]. More than 88% of these glacial microorganisms are likely to be new species with strong spatial and habitat specificity. Also, using the genomes of 883 cultured isolates, the authors constructed the Tibetan Glacier Genome and Gene (TG2G) catalogue, which contains 25 million non-redundant protein-encoding genes, of which 15,954 may be related to functionally novel natural products, only 8.4% of which exist in the current Genome Taxonomy Database (GTDB). The TG2G catalogue is a valuable resource for increasing our understanding of the structure and functions of Tibetan glacial microbiomes, and holds excellent prospects for exploring new biosynthetic potential.

Credit: Philip Patenall/Springer Nature Limited

In another study, Volland et al.[4] found a nearly 1-cm long bacterium, Candidatus *Thiomargarita magnifica*, which is 5,000 times bigger than the average bacterium. *T. magnifica* genomes contain more than 11,000 genes, three times more than the average bacterium. This giant raises questions about additional biochemistry and structural macromolecules that are needed to sustain life, and whether other macro-bacteria exist. Further characterization is needed to elucidate the unique functions and possible biosynthetic capabilities of these bacteria.

The immense phylogenetic, metabolic and functional diversity of Earth's microorganisms provides ample opportunities for discovering new taxa and new natural products with potential biotechnological and therapeutic applications. The question is, how can we leverage microbial bioresources and genetic capabilities to build synthetic biology parts and modules strategically? Owing to biomining and growth of sequencing and informatics tools and the use of machine learning and genome-scale metabolic models, we have begun to unveil the genetic aptitude and potential of microorganisms.

Zhiying Zhao and Yasuo Yoshikuni ✉
The U.S. Department of Energy Joint Genome Institute,
Lawrence Berkeley National Laboratory,
Berkeley, CA, USA.
✉*e-mail: yyoshikuni@lbl.gov*

https://doi.org/10.1038/s41579-022-00795-w

1. Coelho, L. P. et al. Towards the biogeography of prokaryotic genes. *Nature* **601**, 252–256 (2022).
2. Paoli, L. et al. Biosynthetic potential of the global ocean microbiome. *Nature* **607**, 111–118 (2022).
3. Liu, Y. et al. A genome and gene catalog of glacier microbiomes. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-022-01367-2 (2022).
4. Volland, J.-M. et al. A centimeter-long bacterium with DNA contained in metabolically active, membrane-bound organelles. *Science* **376**, 1453–1458 (2022).

Competing interests

The authors declare no competing interests.