

A wealth of new biosynthetic pathways from the global ocean microbiome

DNA from more than 1,000 marine microbial communities around the world was used to reconstruct around 26,000 genomes. The analyses identified a highly biosynthetically diverse family of bacteria in the open ocean, as well as new enzymes and biochemical compounds.

This is a summary of: [Paoli, L. *et al.* Biosynthetic potential of the global ocean microbiome. *Nature* <https://doi.org/10.1038/s41586-022-04862-3> \(2022\).](https://doi.org/10.1038/s41586-022-04862-3)

The mission

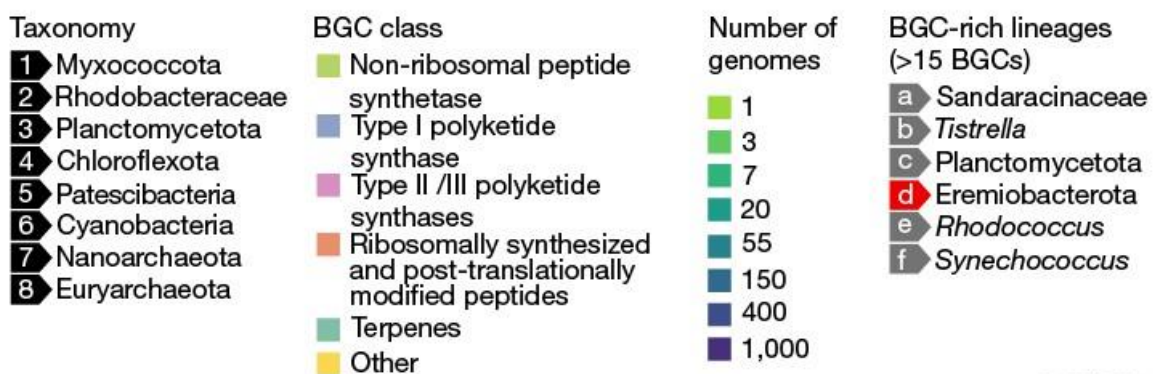
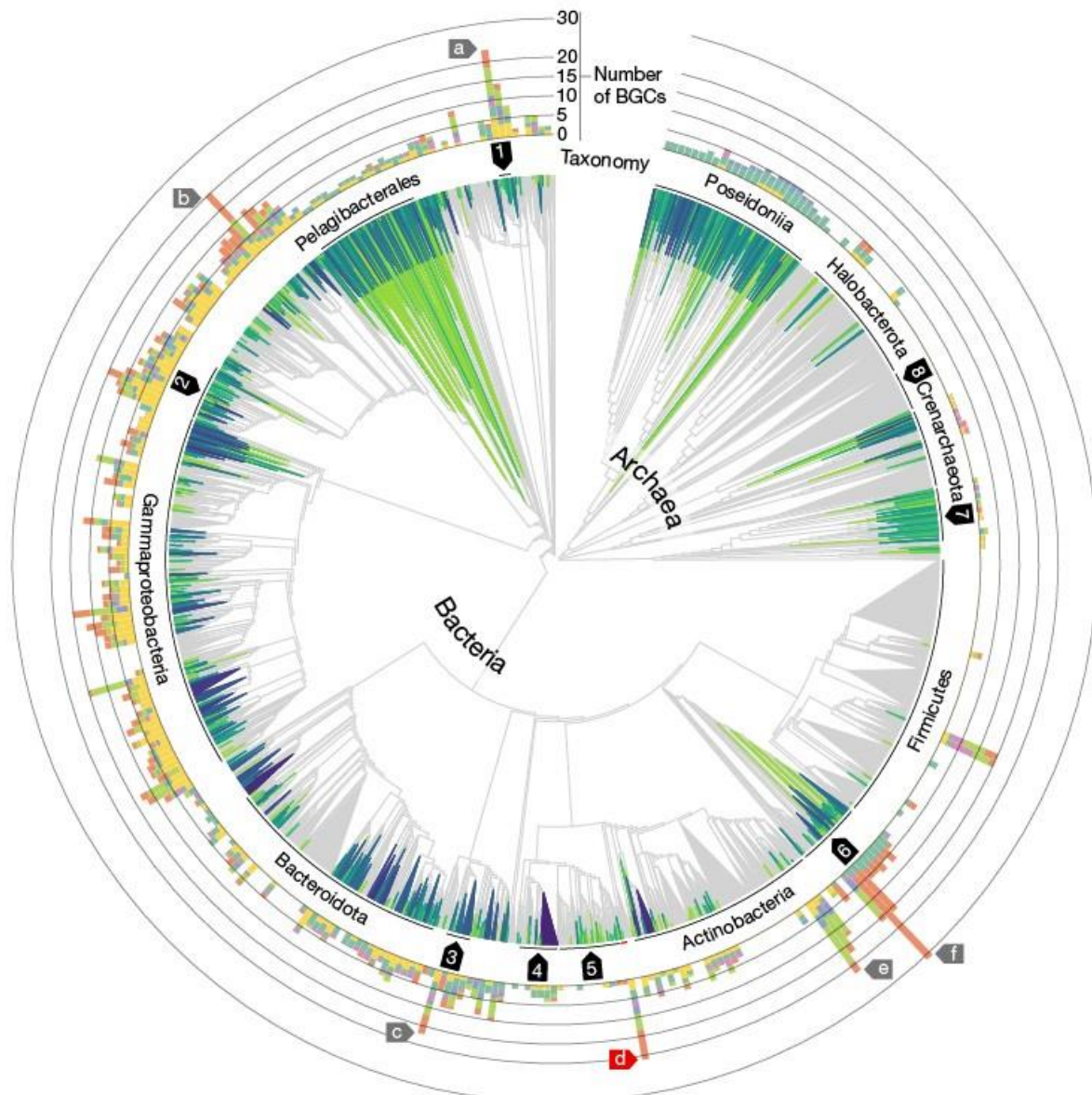
The ocean covers 71% of Earth's surface and contains 97% of its water. Each drop of seawater contains thousands of microbial cells — mostly bacteria and archaea — living in communities that contribute to the global ocean microbiome. However, genetic sequencing of ocean samples has shown that more than two-thirds of marine microbial-community DNA (metagenomes) cannot be associated with known species because most microorganisms cannot currently be grown in the laboratory¹. This means that there is vast, uncharted biodiversity in the global ocean microbiome — particularly in the deep ocean and polar regions²— and probably also many undiscovered biosynthetic pathways that produce biochemical compounds with diverse ecological and cellular functions³. We set out to establish a genomic resource that better represents microbes in oceans around the world and to explore the potential of this chemical treasure trove for the development of useful biocatalysts and therapeutic leads.

The discovery

Thanks to decades of ocean sampling, such as that done by the *Tara* Oceans expedition⁴, we were able to analyse metagenomes from more than 1,000 water samples collected at 215 ocean sites around the world and at several depths. We used state-of-the-art computational methods to reconstruct 26,293 microbial genomes and uncovered more than 2,700 undescribed species. We integrated these with publicly available genomes to build the Ocean Microbiomics Database (Fig. 1), which captured 40–60% of open-ocean metagenomic data. Mining this resource revealed around 40,000 biosynthetic gene clusters (BGCs) — groups of genes predicted to encode pathways for the synthesis of around 7,000 biochemical compounds, more than half of which are likely to be new. In addition to their important ecological roles, many of these compounds or specialized metabolites could have drug-like properties.

This document is the accepted manuscript version of the following article:

Paoli, L., Robinson, S., & Moore, B. (2022). A wealth of new biosynthetic pathways from the global ocean microbiome. *Nature* (2 pp.). <https://doi.org/10.1038/d41586-022-01545-x>



©nature

Figure 1 | Overview of biosynthetic gene clusters in the Ocean Microbiomics

Database. Biosynthetic gene cluster (BGC) diversity overlaid on the phylogenetic tree of a classification system known as the Genome Taxonomy Database⁵. Branches are coloured on the basis of the number of genomes in the Ocean Microbiomics Database for that lineage. Taxonomy is indicated on the tree or using black arrows. Outer bars correspond to the

highest number of BGCs for a given genome in that part of the tree, and are coloured according to the BGC class. Grey arrows indicate BGC-rich lineages. The red arrow indicates the uncultivated phylum 'Ca. Eremiobacterota', which contains the family of bacteria described in this study. Credit: Paoli, L. *et al.*/Nature (CC BY 4.0)

Our analyses uncovered a lineage of bacteria with an unusually high number and diversity of BGCs and showed that it is part of the under-explored phylum 'Candidatus Eremiobacterota'. We named this new family 'Candidatus Eudoremicrobiaceae', after Eudore, the Nereid (sea nymph) of fine gifts in Greek mythology. By expressing two of these BGCs in bacterial hosts, we characterized new biosynthetic enzymes and natural products. The phospeptin pathway produces a peptide that inhibits protease enzymes, providing a chemical template for the development of therapeutics. Furthermore, the pythoamide pathway encodes a massive peptide in a new enzyme family that catalyses the addition of methyl groups to the backbone of peptides — a reaction that could be used in biotechnological processes. These pythoamides are among the most complex natural products revealed so far from uncultivated microbes through synthetic biology.

The implications

Our work provides an unprecedented picture of the biosynthetic activities of microbes in the world's oceans. It generates an atlas to guide future research in fields including marine ecology, evolution, biotechnology and natural-product drug discovery.

We also demonstrated that combined computational and molecular methods can help to identify functions for genetic sequences even for little-known microbial phyla, for which members have not been cultivated in the lab.

We observed some discrepancies between our computational predictions for and biochemical characterization of the phospeptin and pythoamide biosynthetic pathways, showing that there are limits to the accuracy with which we can infer enzyme reactions and natural-product structures by comparing gene sequences with previously known ones. Experimental validation therefore remains an essential step for the discovery of biological functions and compounds.

We chemically characterized only two of thousands of predicted pathways in our data set, meaning that a wealth of biosynthetic enzymes and chemical compounds have yet to be investigated. The Ocean Microbiomics Database is freely available at <https://microbiomics.io/ocean/> as a resource for the research community. We invite scientists worldwide to join us in exploring marine microbial diversity to uncover new enzymes, pathways and specialized metabolites, and to elucidate their ecological functions and potential uses. — **Lucas Paoli is in the Department of**

Biology, Institute of Microbiology and Swiss Institute of Bioinformatics, Swiss Federal Institute of Technology (ETH) Zurich, Zurich, Switzerland, and Serinaa Robinson is at the Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf, Switzerland.

Expert opinion

This paper reports an extensive data analysis of the global ocean microbiome that focuses on the biosynthetic potential of marine bacteria to produce bioactive compounds. The authors identified new bacterial species that are biosynthetically 'gifted' in terms of encoding diverse and plentiful biosynthetic gene clusters, and provide experimental data to substantiate their claims. The sheer volume of data presented and the identification of a new biosynthetically rich marine lineage is impressive. — **Bradley Moore is at the University of California, San Diego, San Diego, California, USA.**

Behind the paper

For the past two decades, scientists around the globe have been sampling and sequencing the DNA of marine microbes as part of expeditions and surveys such as *Tara* Oceans, BioGEOTRACES, Malaspina, HOT and BATS, or through projects such as the Global Ocean Reference Genomes and Mar databases. This project and collaboration is built entirely on that work. The study took a crucial turn in the social room of the Institute of Microbiology at ETH Zurich in 2019. As the coffee machine purred, J. Piel suggested, "You know what we could do with all these genomes? We could look for BGCs, that would be very exciting." The resulting project soon led to another of many memorable moments when three people at a computer screen contemplated the genome of a new bacterium — what would become '*Ca.Eudoremicrobium malaspinii*'. We are overjoyed to deliver this resource back to the research community and hope that the Ocean Microbiomics Database will deepen the understanding of the world's waters for decades to come. — **L.P.**