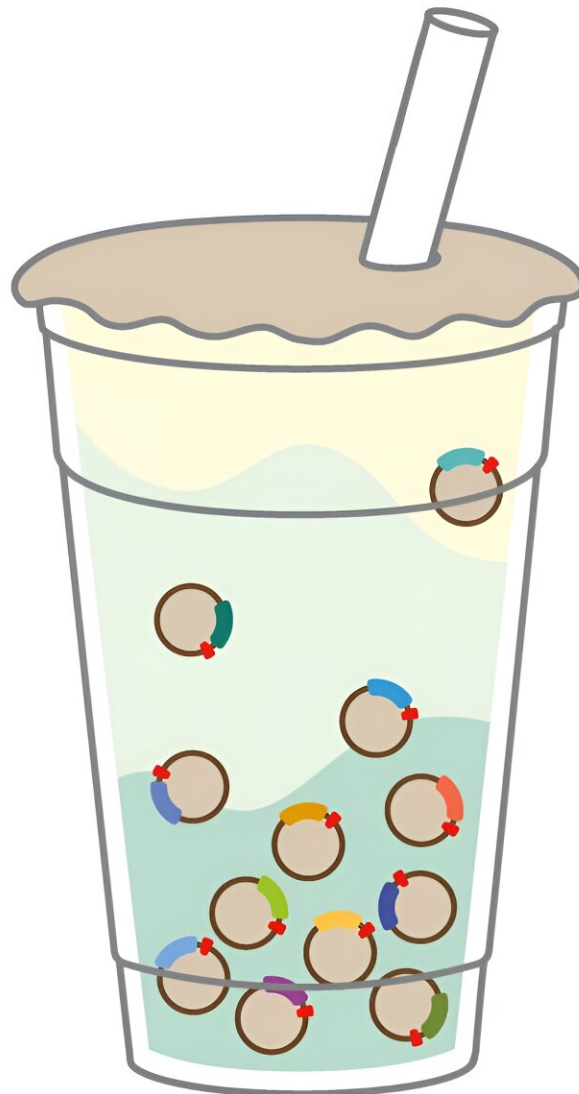


Revealing the mysteries within microbial genomes with a new high-throughput approach

August 14 2024, by Aliyah Kovner



The Boba-seq logo, created by Davian Ho, a former undergraduate student researcher in the Arkin lab. Credit: Davian Ho

A new technique developed at Lawrence Berkeley National Laboratory (Berkeley Lab) will make it much easier for researchers to discover the traits or activities encoded by genes of unknown function in microbes, a key step toward understanding the roles and impact of individual species.

The approach, called barcoded overexpression bacterial shotgun library sequencing, or Boba-seq, is described in a paper [published](#) August 5 in *Nature Communications*.

"There is so much genetic dark matter—DNA that we can sequence quickly with today's methods but don't know the function of—out there in the microbial universe. And the question is, how are we ever going to study all that matter to understand the microbiomes surrounding us? The fundamental answer is—like this," said senior author Adam Arkin, a senior faculty scientist in Berkeley Lab's Biosciences Area.

Boba-seq involves taking random fragments of DNA from bacteria of interest and expressing them in host [bacterial cells](#).

"The idea is that we can see how the presence of new genes confer a difference in phenotype of growth of that bacterium," said first author Yolanda Huang, an assistant professor of Microbiology and Immunology at the University of Buffalo who was a postdoctoral researcher in Arkin's lab at the time of the study. "This is a [functional genomics](#) approach that we can use to quickly connect a gene or a piece of DNA sequence with a potential function."

The term "barcode" in the name refers to a small sequence of DNA that the scientists use as an identifying tag for a much larger fragment of DNA, much like how a barcode at a grocery store identifies a specific item with a small code. The entire genome of the organism being studied is randomly separated into fragments containing single genes or clusters of several genes, then inserted into plasmids—circular packages of DNA—that have been tagged with unique barcodes.

The Boba-seq "library" refers to all the barcoded plasmids containing fragments from an organism. This library can be introduced into different bacterial hosts to generate a huge number of genetic variants, which are then screened for new behaviors or properties.

Arkin and his colleagues in the Biosciences Area are leading experts on high-throughput techniques to study gene function, and have been involved in the invention of a number of other techniques that either insert or silence genes to investigate their function, including RB-TnSeq, CRISPRi and [Dub-Seq](#).

With Boba-seq, hundreds of thousands of barcoded fragments can be put into host cells and cultured under varying conditions to determine function in a single experiment. For example, if cells with a certain barcode grow happily when the whole culture is exposed to an antibiotic, but the others perish, you know that the gene or genes in that fragment encoded antibiotic resistance traits. And identifying the fragment responsible for this new ability is cheap and fast, thanks to the barcode.

"Yolanda's innovations with Boba-seq allows us to identify which of hundreds of thousands of fragments are conferring the phenotype or the property that we want," said Arkin. "Our new strategy allows us to make libraries and use them at a higher throughput than previous overexpression approaches."

The other significant breakthrough is that Boba-seq fragments can be tested in the same organism that they were pulled from (or a close relative), which is essential for getting an accurate picture of what a gene does. Previous techniques were limited because they only tested genes inside model organisms like *E. coli* and yeast. Genes from organisms very different from *E. coli* are often not functional in *E. coli*, making it difficult or impossible to get a clear picture of what the genes do.

The computational tool used to process results from the laboratory work involved in Boba-seq is available to other researchers on an open-source platform.

"I am excited to see how others may use Boba-seq worldwide, especially for metagenomic studies of the gut or environment," said co-author Allison Hung, a UC Berkeley graduate student in the Arkin lab. "The ability to extract functional information from a microbial community without isolation saves a huge amount of time and resources, and will be key for studying microbes that are difficult to culture in a lab, like those living in complex ecosystems currently studied as part of ENIGMA."

ENIGMA, short for Ecosystems and Networks Integrated with Genes and Molecular Assemblies, is a Department of Energy (DOE) Scientific Focus Area co-led by Arkin that is aimed at understanding how microbial communities cycle nutrients through ecosystems and detoxify toxic heavy metal contaminants.

After building and refining Boba-seq, Arkin's team tested the new technique by studying the genes in Bacteroidales, a taxonomic order of microbes that are abundant in the human gut and known to play many roles in our internal microbiome. Bacteroidales are also major players in terrestrial soil processes, where they degrade organic matter and return the nutrients to plants. The team generated 305,000 barcoded fragments from libraries of six Bacteroidales species and evaluated more than

21,000 protein-coding genes in parallel.

Results from these proof-of-principle experiments revealed that genes encoding enzymes that build certain lipid molecules endow resistance to ceftriaxone, an antibiotic in the cephalosporin class. These genes have not been previously linked to antibiotic resistance, and warrant further investigation.

The team also discovered several new functions in carbohydrate metabolism, including an enzyme needed to metabolize glucosamine, a modified sugar molecule found in bones, connective tissue, and the exoskeletons of insects and crustaceans. In the gut, microbes use glucosamine as an energy molecule and to construct their cell walls, whereas human cells that form the lining of the intestine use it to produce the mucus membrane that helps maintain healthy nutrient uptake and prevent invasion of pathogens.

These insights into Bacteroidales will help health researchers better understand gut function, as this order acts as "commensals most of the time and really maintain gut health," said Huang. "But in certain states, the nutrient released by Bacteroidales can be used by pathogens to support their own growth."

Arkin and his ENIGMA colleagues are now using Boba-seq to investigate how soil microbes derive energy from complex carbon-based molecules in the environment that most life forms cannot metabolize. Meanwhile, Huang plans to use Boba-seq in her new lab at the University of Buffalo's Jacobs School of Medicine & Biomedical Sciences to study [genes](#) that bacteria use to evade attacks from bacteriophages (viruses that infect bacteria), increase colonization efficiency in the gut, and break down complex carbohydrates.

More information: Yolanda Y. Huang et al, Barcoded overexpression

screens in gut Bacteroidales identify genes with roles in carbon utilization and stress resistance, *Nature Communications* (2024). [DOI: 10.1038/s41467-024-50124-3](https://doi.org/10.1038/s41467-024-50124-3)

Provided by Lawrence Berkeley National Laboratory

Citation: Revealing the mysteries within microbial genomes with a new high-throughput approach (2024, August 14) retrieved 14 August 2024 from <https://phys.org/news/2024-08-revealing-mysteries-microbial-genomes-high.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.