MICROBIOLOGY

Exploring the chemical space of the human microbiome

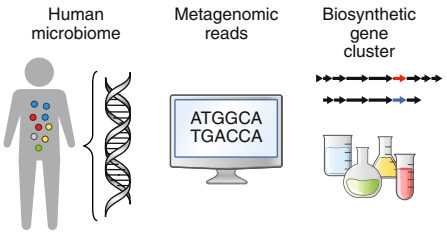
An in silico approach helps researchers identify biosynthetic gene clusters coding for bioactive small molecules from metagenomic data of the human microbiome.

It is now well appreciated that the microbiome has an important role in human health and disease, yet the underlying molecule-mediated mechanisms governing microbiome–host interactions are largely unexplored. "A key route towards defining these mechanisms is through the identification of biologically active small molecules produced by the microbiome," says Mohamed Donia, a chemical biologist from Princeton University.

The chemical capabilities of microbiome-derived small molecules can influence biological interactions by directly targeting human cells or receptors or by indirectly mediating other members existing in the microbiota.

Mohamed Donia has been working on small molecules from the human microbiome and their roles since 2010, when he was a postdoctoral fellow with Michael Fischbach, then at the University of California, San Francisco (now at Stanford University). Donia noted that the genetics of small-molecule production, i.e., biosynthetic gene clusters (BGCs), can serve as an initial proxy for small molecule discovery. Therefore, he and colleagues pinpointed the gene clusters by analyzing genomes of bacteria isolated from the human microbiome, in which they identified more than 3,000 BGCs and revealed BGCs for a class of antibiotics that are widely distributed in the human microbiome.

"There are several great tools to identify BGCs from bacterial genomes, including ones that we use every day in my lab [antiSMASH]," he says, "but the main requirement for these tools is to have an assembled bacterial genome or at least a fully assembled BGC in the first place." This type of tool is thus limited by the facts that the majority of species in the human microbiome space have not yet been isolated or cultivated and that most isolation efforts have concentrated on the healthy gut microbiome from Western populations, which may not represent microbiomes found in disease states, other body sites or non-Western populations. At the same time, assemblies of complex metagenomic sequencing data often yield fragmented BGCs and are biased towards the most abundant members of the microbiome.



Detecting biosynthetic gene clusters from metagenomic reads. Credit: Marina Corral Spence/Springer Nature

Donia and his colleagues now move the goal to developing algorithms for identifying BGCs at the single-metagenomic read level, "from amongst billions of reads and without the need for bacterial isolation, reference genomes or initial metagenomics assemblies," Donia says.

Here is where MetaBGC comes in. MetaBGC is a sensitive algorithm relying on repurposing established probabilistic models, profiling hidden Markov models to detect metagenomic reads that encode the protein family of interest in complex human microbiome datasets. The ability to detect BGCs in a de novo manner offers an opportunity to map microbiome-derived small molecules using public metagenomics datasets.

Moreover, detecting BGCs directly from metagenomics datasets circumvents cultivation difficulties, as one can detect BGCs from sequenced bacteria that have been cultured or from yet-uncultivated strains. The advantage becomes apparent when the bacteria of interest are in low abundance or poorly represented in metagenomics assemblies.

Using MetaBGC, Sugimoto and Camacho, Donia's postdoctoral fellow and PhD student, respectively, focused on a class of BGCs that were underexplored in the human microbiome, type II polyketide synthase (TII-PKS) BGCs. These BGCs often encode small molecules with biological functions, such as anticancer and antibiotic drugs.

"Strikingly, we discovered 13 complete gene clusters [encoding TII-PKS], which are widely distributed in the gut, oral, and skin microbiome of people all the way from the US to Fiji," Donia said. They further heterologously expressed these identified BGCs and directly characterized their chemical products. They verified two of these gene clusters, including one from the mouth and one from the gut, and showed that two of the five discovered molecules are potent antibiotics, like their clinically used relatives. In antimicrobial assays, the molecules identified from the oral microbiome yielded antibacterial activities against neighboring microbes, suggesting a possible mechanism of niche competition or defense against pathogens.

Donia said, "This is the beginning of this effort. We still don't know any biologically activity for three of the five molecules we found, and we don't have molecules from 11 of the 13 clusters we discovered." With the tools available for discovering BGCs encoding for interesting molecules, researchers can start to explore a chemical space that has direct implications for the mechanistic understanding of microbiome–host interactions.

Lei Tang

Published online: 28 November 2019
https://doi.org/10.1038/s41592-019-0671-9

Research paper
Sugimoto, Y. et al. A metagenomic strategy for harnessing the chemical repertoire of the human microbiome. *Science* https://doi.org/10.1126/science.aax9176 (2019).