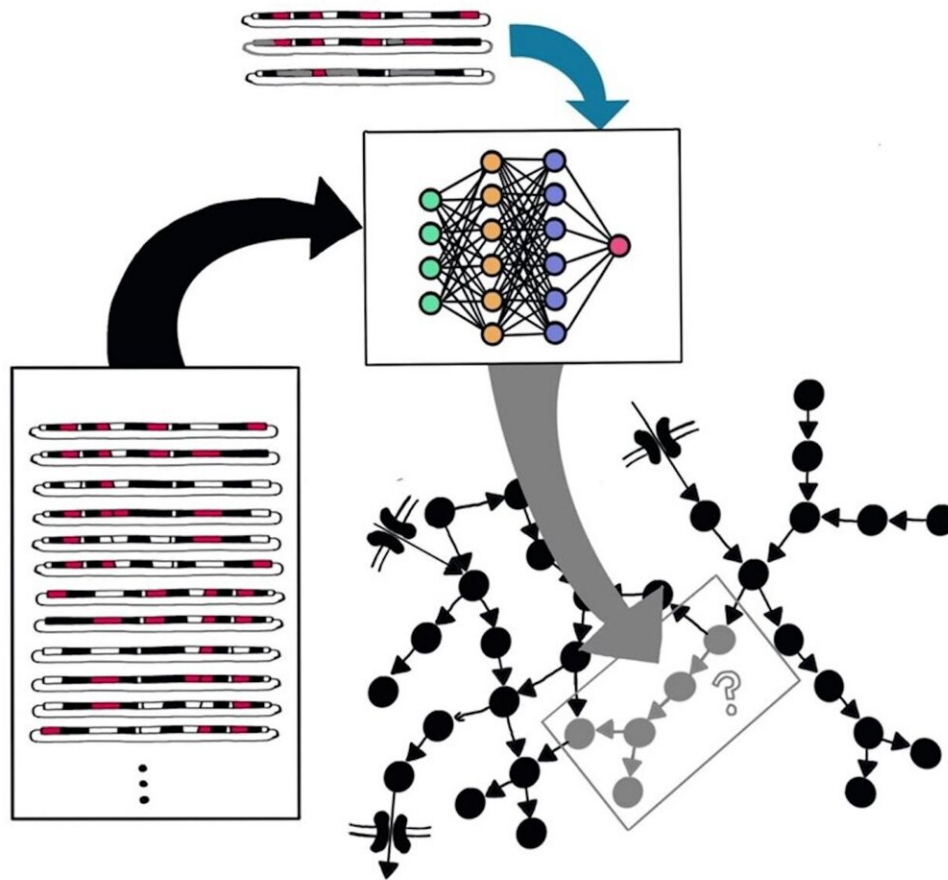


Predicting metabolic potential in bacteria from limited genome data

August 8 2024



MetaPathPredict is a novel algorithm that uses machine learning to predict metabolic pathways from limited information. Credit: Jason McDermott/Pacific Northwest National Laboratory

How bacteria eat food, and what kinds of products they can make from that food, is dictated by the metabolic network of enzyme patterns encoded in their genomes. Using computational methods to learn these patterns across a large number of known bacteria allows the genome of a new bacteria to be analyzed. This reveals what kind of metabolism it is capable of—even when only partial information is provided, which is common in environmental samples.

This project was initiated by David Geller-McGrath as part of his graduate thesis project at the Woods Hole Oceanographic Institute under Dr. Edgcomb. He refined the approaches and developed the code during his time as an Office of Science Graduate Student Research fellowship in computational biosciences working with Dr. McDermott at Pacific Northwest National Laboratory, and with Dr. Wheeler at the University of Arizona.

This new computational method, now [published](#) in *eLife*, enables the discovery of new metabolic capabilities for bacteria important for the environment and bioenergy applications. This is important for understanding microbiomes (communities of bacteria and other microorganisms) that support plant growth for improved [crop yields](#). In addition, a better understanding of different metabolic networks will allow new ways of engineering bacteria for other bioenergy and biomedical applications.

The method learns patterns of proteins present in [metabolic pathways](#) from a large collection of annotated bacterial genomes using a deep learning approach. A significant advantage of this tool is that it is designed and tested on incomplete genomic data. This allows [bacterial genomes](#) to be identified and assessed for metabolic potential in complex microbiomes from soil or other sources, samples that are often incomplete.

More information: David Geller-McGrath et al, Predicting metabolic modules in incomplete bacterial genomes with MetaPathPredict, *eLife* (2024). [DOI: 10.7554/eLife.85749](https://doi.org/10.7554/eLife.85749)

Provided by Pacific Northwest National Laboratory

Citation: Predicting metabolic potential in bacteria from limited genome data (2024, August 8) retrieved 10 August 2024 from <https://phys.org/news/2024-08-metabolic-potential-bacteria-limited-genome.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.