

New machine learning approach could accelerate bioengineering

May 30 2018, by Dan Krotz



A new approach developed by Zak Costello (left) and Hector Garcia Martin brings the the speed and analytic power of machine learning to bioengineering. Credit: Marilyn Chung, Berkeley Lab

Scientists from the Department of Energy's Lawrence Berkeley National

Laboratory (Berkeley Lab) have developed a way to use machine learning to dramatically accelerate the design of microbes that produce biofuel.

Their computer algorithm starts with abundant data about the proteins and metabolites in a biofuel-producing microbial [pathway](#), but no information about how the pathway actually works. It then uses data from previous experiments to learn how the pathway will behave. The scientists used the technique to automatically predict the amount of biofuel produced by pathways that have been added to E. coli bacterial cells.

The new approach is much faster than the current way to predict the behavior of pathways, and promises to speed up the development of biomolecules for many applications in addition to commercially viable biofuels, such as drugs that fight antibiotic-resistant infections and crops that withstand drought.

The research was published May 29 in the journal *Nature Systems Biology and Applications*.

In biology, a pathway is a series of chemical reactions in a cell that produce a specific compound. Researchers are exploring ways to re-engineer pathways, and import them from one microbe to another, to harness nature's toolkit to improve medicine, energy, manufacturing, and agriculture. And thanks to new synthetic biology capabilities, such as the gene-editing tool CRISPR-Cas9, scientists can conduct this research at a precision like never before.

"But there's a significant bottleneck in the development process," said Hector Garcia Martin, group lead at the DOE Agile BioFoundry and director of Quantitative Metabolic Modeling at the Joint BioEnergy Institute (JBEI), a DOE Bioenergy Research Center funded by DOE's

Office of Science and led by Berkeley Lab. The research was performed by Zak Costello (also with the Agile BioFoundry and JBEI) under the direction of Garcia Martin. Both researchers are also in Berkeley Lab's Biological Systems and Engineering Division.

"It's very difficult to predict how a pathway will behave when it's re-engineered. Trouble-shooting takes up 99% of our time. Our approach could significantly shorten this step and become a new way to guide bioengineering efforts," Garcia Martin added.

The current way to predict a pathway's dynamics requires a maze of differential equations that describe how the components in the system change over time. Subject-area experts develop these "kinetic models" over several months, and the resulting predictions don't always match experimental results.

Machine learning, however, uses data to train a computer algorithm to make predictions. The algorithm learns a system's behavior by analyzing data from related systems. This allows scientists to quickly predict the function of a pathway even if its mechanisms are poorly understood—as long as there are enough data to work with.

The scientists tested their technique on pathways added to *E. coli* cells. One pathway is designed to produce a bio-based jet fuel called limonene; the other produces a gasoline replacement called isopentenol. Previous experiments at JBEI yielded a trove of data related to how different versions of the pathways function in various *E. coli* strains. Some of the strains have a pathway that produces small amounts of either limonene or isopentenol, while other strains have a version that produces large amounts of the biofuels.

The researchers fed this data into their algorithm. Then machine learning took over: The algorithm taught itself how the concentrations of

metabolites in these pathways change over time, and how much biofuel the pathways produce. It learned these dynamics by analyzing data from the two experimentally known pathways that produce small and large amounts of biofuels.

The algorithm used this knowledge to predict the behavior of a third set of "mystery" pathways the algorithm had never seen before. It accurately predicted the biofuel-production profiles for the mystery pathways, including that the pathways produce a medium amount of fuel. In addition, the machine learning-derived prediction outperformed kinetic models.

"And the more data we added, the more accurate the predictions became," said Garcia Martin. "This approach could expedite the time it takes to design new biomolecules. A project that today takes ten years and a team of experts could someday be handled by a summer student."

More information: Zak Costello et al. A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data, *npj Systems Biology and Applications* (2018). [DOI: 10.1038/s41540-018-0054-3](https://doi.org/10.1038/s41540-018-0054-3)

Provided by Lawrence Berkeley National Laboratory

Citation: New machine learning approach could accelerate bioengineering (2018, May 30) retrieved 20 September 2024 from <https://phys.org/news/2018-05-machine-approach-bioengineering.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.