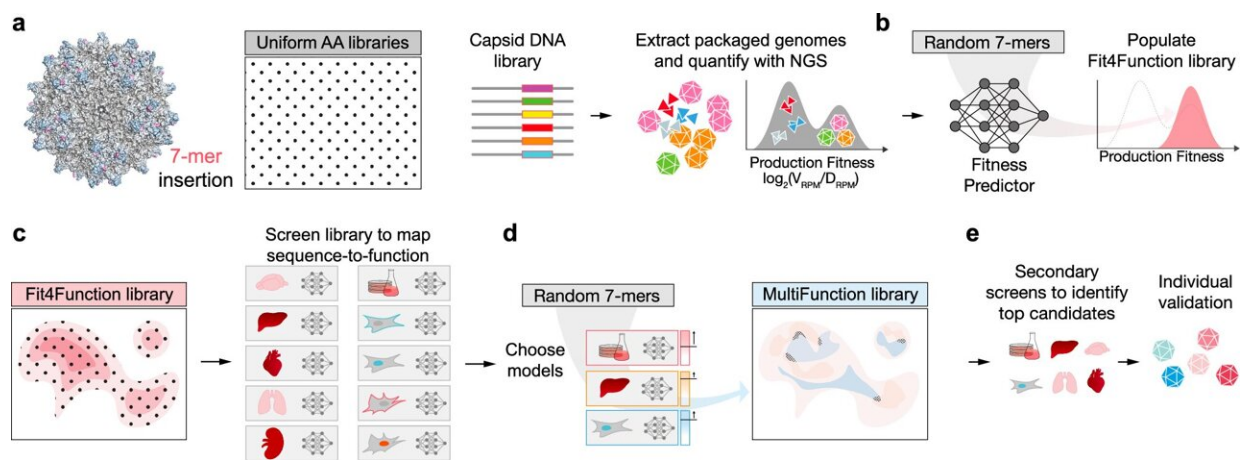


Machine learning approach helps researchers design better gene-delivery vehicles for gene therapy

August 8 2024, by Alessandra DiCorato



Systematic multi-trait protein optimization paradigm. Credit: *Nature Communications* (2024). DOI: 10.1038/s41467-024-50555-y

Gene therapy could potentially cure genetic diseases, but it remains a challenge to package and deliver new genes to specific cells safely and effectively. Existing methods of engineering one of the most commonly used gene-delivery vehicles, adeno-associated viruses (AAV), are often slow and inefficient.

Now, researchers at the Broad Institute of MIT and Harvard have developed a machine-learning approach that promises to speed up AAV

engineering for [gene therapy](#). The tool helps researchers engineer the protein shells of AAVs, called capsids, to have multiple desirable traits, such as the ability to deliver cargo to a specific organ but not others or to work in multiple species. Other methods only look for capsids that have one trait at a time.

The team used their approach to design capsids for a commonly used type of AAV called AAV9 that more efficiently targeted the liver and could be easily manufactured.

They found that about 90% of the capsids predicted by their machine learning models successfully delivered their cargo to [human liver cells](#) and met five other key criteria.

They also found that their machine learning model correctly predicted the behavior of the proteins in [macaque monkeys](#) even though it was trained only on mouse and human cell data. This finding suggests that the new method could help scientists more quickly design AAVs that work across species, which is essential for translating gene therapies to humans.

[The findings](#), which appear in *Nature Communications*, come from the lab of Ben Deverman, institute scientist and director of vector engineering at the Stanley Center for Psychiatric Research at the Broad. Fatma-Elzahraa Eid, a senior machine learning scientist in Deverman's group, was the first author on the study.

"This was a really unique approach," Deverman said. "It highlights the importance of wet lab biologists working with machine learning scientists early to design experiments that generate machine learning enabling data rather than as an afterthought."

Group leader Ken Chan, graduate student Albert Chen, research

associate Isabelle Tobey, and scientific advisor Alina Chan, all in Deverman's lab, also contributed significantly to the study.

Make way for machines

Traditional approaches for designing AAVs involve generating large libraries containing millions of capsid protein variants and then testing them in cells and animals in several rounds of selection. This process can be costly and time-consuming, and generally results in researchers identifying only a handful of capsids that have a specific trait. This makes it challenging to find capsids that meet multiple criteria.

Other groups have used machine learning to expedite large-scale analysis, but most methods optimized proteins for one function at the expense of another.

Deverman and Eid realized that datasets based on existing large AAV libraries weren't well suited for training machine learning models. "Instead of just taking data and giving it to machine learning scientists, we thought, 'What do we need to train machine learning models better?'" Eid said. "Figuring that out was really instrumental."

They first used an initial round of machine learning modeling to generate a new moderately sized library, called Fit4Function, that contained capsids that were predicted to package gene cargo well. The team screened the library in human cells and mice to find capsids that had specific functions important for gene therapy in each species.

They then used that data to build multiple machine learning models that could each predict a certain function from a [capsid](#)'s amino acid sequence. Finally, they used the models in combination to create "multifunction" libraries of AAVs optimized for multiple traits at once.

The future of protein design

As proof of concept, Eid and other researchers in Deverman's lab combined six models to design a library of capsids that had multiple desired functions, including manufacturability and the ability to target the liver across human cells and mice. Almost 90% of these proteins displayed all of the desired functions simultaneously.

The researchers also found that the model—trained only on data from mice and human cells—correctly predicted how AAVs are distributed to different organs of macaques, suggesting that these AAVs do this through a mechanism that translates across species. That could mean that in the future, gene therapy researchers could more quickly identify capsids with multiple desirable properties for human use.

In the future, Eid and Deverman say their models could help other groups create gene therapies that either target or specifically avoid the liver. They also hope that other labs will use their approach to generate models and libraries of their own that, together, could form a machine-learning atlas: a resource that could predict the performance of AAV capsids across dozens of traits to accelerate gene therapy development.

More information: Fatma-Elzahraa Eid et al, Systematic multi-trait AAV capsid engineering for efficient gene delivery, *Nature Communications* (2024). [DOI: 10.1038/s41467-024-50555-y](https://doi.org/10.1038/s41467-024-50555-y)

Provided by Broad Institute of MIT and Harvard

Citation: Machine learning approach helps researchers design better gene-delivery vehicles for gene therapy (2024, August 8) retrieved 8 August 2024 from <https://medicalxpress.com/news/2024-08-machine-approach-gene-delivery-vehicles.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.