AI combined with CRISPR precisely controls gene expression

July 3 2023

Credit: Unsplash/CC0 Public Domain

Artificial intelligence can predict on- and off-target activity of CRISPR tools that target RNA instead of DNA, according to new research published in *Nature Biotechnology*.

The study by researchers at New York University, Columbia University, and the New York Genome Center, combines a deep learning model

with CRISPR screens to control the expression of human [genes](link) in different ways—such as flicking a light switch to shut them off completely or by using a dimmer knob to partially turn down their activity. These precise gene controls could be used to develop new CRISPR-based therapies.

CRISPR is a gene editing technology with many uses in biomedicine and beyond, from [treating sickle cell anemia](link) to engineering tastier mustard greens. It often works by targeting DNA using an enzyme called Cas9. In recent years, scientists discovered another type of CRISPR that instead targets RNA using an enzyme called Cas13.

RNA-targeting CRISPRs can be used in a wide range of applications, including RNA editing, knocking down RNA to block expression of a particular gene, and high-throughput screening to determine promising drug candidates. Researchers at NYU and the New York Genome Center created [a platform for RNA-targeting CRISPR screens using Cas13](link) to better understand RNA regulation and to identify the function of non-coding RNAs. Because RNA is the main genetic material in viruses including SARS-CoV-2 and flu, [RNA-targeting CRISPRs also hold promise for developing new methods to prevent or treat viral infections](link). Also, in human cells, when a gene is expressed, one of the first steps is the creation of RNA from the DNA in the genome.

A key goal of the study is to maximize the activity of RNA-targeting CRISPRs on the intended target RNA and minimize activity on other RNAs which could have detrimental side effects for the cell. Off-target activity includes both mismatches between the guide and target RNA as well as insertion and deletion mutations.

Earlier studies of RNA-targeting CRISPRs focused only on on-target activity and mismatches; predicting off-target activity, particularly insertion and deletion mutations, has not been well-studied. In human

populations, about one in five mutations are insertions or deletions, so these are important types of potential off-targets to consider for CRISPR design.

"Similar to DNA-targeting CRISPRs such as Cas9, we anticipate that RNA-targeting CRISPRs such as Cas13 will have an outsized impact in molecular biology and biomedical applications in the coming years," said Neville Sanjana, associate professor of biology at NYU, associate professor of neuroscience and physiology at NYU Grossman School of Medicine, a core faculty member at New York Genome Center, and the study's co-senior author. "Accurate guide prediction and off-target identification will be of immense value for this newly developing field and therapeutics."

In their study in *Nature Biotechnology*, Sanjana and his colleagues performed a series of pooled RNA-targeting CRISPR screens in human cells. They measured the activity of 200,000 guide RNAs targeting essential genes in human cells, including both "perfect match" guide RNAs and off-target mismatches, insertions, and deletions.

Sanjana's lab teamed up with the lab of machine learning expert David Knowles to engineer a deep learning model they named TIGER (Targeted Inhibition of Gene Expression via guide RNA design) that was trained on the data from the CRISPR screens. Comparing the predictions generated by the deep learning model and laboratory tests in human cells , TIGER was able to predict both on-target and off-target activity, outperforming previous models developed for Cas13 on-target guide design and providing the first tool for predicting off-target activity of RNA-targeting CRISPRs.

"Machine learning and deep learning are showing their strength in genomics because they can take advantage of the huge datasets that can now be generated by modern high-throughput experiments. Importantly,

we were also able to use 'interpretable machine learning' to understand why the model predicts that a specific guide will work well," said Knowles, assistant professor of computer science and systems biology at Columbia University's School of Engineering and Applied Science, a core faculty member at New York Genome Center, and the study's co-senior author.

"Our earlier research demonstrated how to design Cas13 guides that can knock down a particular RNA. With TIGER, we can now design Cas13 guides that strike a balance between on-target knockdown and avoiding off-target activity," said Hans-Hermann (Harm) Wessels, the study's co-first author and a senior scientist at the New York Genome Center, who was previously a postdoctoral fellow in Sanjana's laboratory.

The researchers also demonstrated that TIGER's off-target predictions can be used to precisely modulate gene dosage—the amount of a particular gene that is expressed—by enabling partial inhibition of gene expression in cells with mismatch guides. This may be useful for diseases in which there are too many copies of a gene, such as Down syndrome, certain forms of schizophrenia, Charcot-Marie-Tooth disease (a hereditary nerve disorder), or in cancers where aberrant gene expression can lead to uncontrolled tumor growth.

"Our deep learning model can tell us not only how to design a guide RNA that knocks down a transcript completely, but can also 'tune' it—for instance, having it produce only 70% of the transcript of a specific gene," said Andrew Stirn, a Ph.D. student at Columbia Engineering and the New York Genome Center, and the study's co-first author.

By combining artificial intelligence with an RNA-targeting CRISPR screen, the researchers envision that TIGER's predictions will help avoid undesired off-target CRISPR activity and further spur development of a

new generation of RNA-targeting therapies.

"As we collect larger datasets from CRISPR screens, the opportunities to apply sophisticated machine learning models are growingly rapid. We are lucky to have David's lab next door to ours to facilitate this wonderful, cross-disciplinary collaboration. And, with TIGER, we can predict off-targets and precisely modulate gene dosage which enables many exciting new applications for RNA-targeting CRISPRs for biomedicine," said Sanjana.

This latest study further advances the broad applicability of RNA-targeting CRISPRs for human genetics and drug discovery, building on the NYU team's prior work to develop guide RNA design rules, target RNAs in diverse organisms including viruses like SARS-CoV-2, engineer protein and RNA therapeutics, and leverage single-cell biology to reveal synergistic drug combinations for leukemia.

More information: Prediction of on-target and off-target activity of CRISPR–Cas13d guide RNAs using deep learning, *Nature Biotechnology* (2023). DOI: 10.1038/s41587-023-01830-8

Provided by New York University