

New tool enables scientists to interpret 'dark matter' DNA

April 4 2016



A depiction of the double helical structure of DNA. Its four coding units (A, T, C, G) are color-coded in pink, orange, purple and yellow. Credit: NHGRI

Scientists at the Gladstone Institutes have invented a new way to read and interpret the human genome. The computational method, called TargetFinder, can predict where non-coding DNA—the DNA that does not code for proteins—interacts with genes. This technology helps researchers connect mutations in the so-called genomic "dark matter" with the genes they affect, potentially revealing new therapeutic targets for genetic disorders.

In the study, published in *Nature Genetics*, the researchers looked at fragments of non-coding DNA called [enhancers](#). Enhancers act like an instruction manual for a gene, dictating when and where a gene is turned on. Genes can be separated from their enhancers by long stretches of DNA that contain many other [genes](#).

"Most genetic mutations that are associated with disease occur in enhancers, making them an incredibly important area of study," said senior author Katherine Pollard, PhD, a senior investigator at the Gladstone Institutes. "Before now, we struggled to understand how enhancers find the distant genes they act upon."

Scientists originally believed that enhancers mostly affect the gene nearest to them. However, the new study revealed that, on a strand of DNA, enhancers can be millions of letters away from the gene they influence, skipping over the genes in between. When an enhancer is far away from the gene it affects, the two connect by forming a three-dimensional loop, like a bow on the genome.

Using machine learning technology, the researchers analyzed hundreds of existing datasets from six different cell types to look for patterns in the genome that identify where a gene and enhancer interact. They discovered several patterns that exist on the loops that connect enhancers to genes. This pattern accurately predicted whether a gene-enhancer interaction occurred 85 percent of the time.

"It's remarkable that we can predict complex three-dimensional interactions from relatively simple data," said first author Sean Whalen, PhD, a biostatistician at Gladstone. "No one had looked at the information stored on loops before, and we were surprised to discover how important that information is."

Performing experiments in the lab to identify all of these gene-enhancer interactions can take millions of dollars and years of research. The new computational approach is a much cheaper and less time-consuming way to identify gene-enhancer connections in the genome. The technology also provides insight into how DNA loops form and how they might break in disease. The scientists have offered all of the code and data from TargetFinder online for free.

"Our ability to predict the gene targets of enhancers so accurately enables us to link mutations in enhancers to the genes they target," said Pollard. "Having that link is the first step towards using these connections to treat diseases."

More information: Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin, *Nature Genetics*, [DOI: 10.1038/ng.3539](https://doi.org/10.1038/ng.3539)

Provided by Gladstone Institutes

Citation: New tool enables scientists to interpret 'dark matter' DNA (2016, April 4) retrieved 19 April 2024 from <https://medicalxpress.com/news/2016-04-tool-enables-scientists-dark-dna.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.