

CRISPR and single-cell sequencing pinpoint causal genetic variants for traits and diseases

May 4 2023



Credit: CC0 Public Domain

A major challenge in human genetics is understanding which parts of the genome drive specific traits or contribute to disease risk. This challenge is even greater for genetic variants found in the 98% of the genome that does not encode proteins.

A new approach developed by researchers at New York University and the New York Genome Center combines genetic association studies, [gene editing](#), and single-cell sequencing to address these challenges and discover causal variants and genetic mechanisms for blood cell traits.

Their approach, dubbed STING-seq and published in *Science*, addresses the challenge of directly connecting genetic variants to human traits and health, and can help scientists identify [drug targets](#) for diseases with a genetic basis.

Over the past two decades, [genome-wide association studies](#) (GWAS) have become an important tool for studying the human genome. Using GWAS, scientists have identified thousands of genetic mutations or variants associated with many diseases, from schizophrenia to diabetes, as well as traits such as height. These studies are conducted by comparing the genomes of large populations to find variants that occur more often in those with a specific disease or trait.

GWAS can reveal what regions of the genome and potential variants are implicated in diseases or traits. However, these associations are nearly always found in the 98% of the genome that does not code for proteins, which is much less well understood than the well-studied 2% of the genome that codes for proteins. A further complication is that many variants are found in close proximity to each other within the genome

and travel together through generations, a concept known as linkage. This can make it difficult to tease apart which variant plays a truly causal role from other variants that are just located nearby. Even when scientists can identify which variant is causing a disease or trait, they do not always know what gene the variant impacts.

"A major goal for the study of human diseases is to identify causal genes and variants, which can clarify biological mechanisms and inform drug targets for these diseases," 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.

"The huge success in GWAS has highlighted the challenge of extracting insights into disease biology from these massive data sets. Despite all of our efforts during the past 10 years, the glass was still just half full—at best. We needed a new approach," said Tuuli Lappalainen, senior associate faculty member at the New York Genome Center, professor of genomics at the KTH Royal Institute of Technology in Sweden, and the study's co-senior author.

A cure for sickle cell anemia

A recent scientific breakthrough in the treatment of [sickle cell anemia](#)—a genetic disorder marked by episodes of intense pain—illustrates how combining GWAS with cutting-edge molecular tools like gene editing can identify causal variants and lead to innovative therapies. Using GWAS, scientists identified areas of the genome important for producing [fetal hemoglobin](#), a target based on its promise for reversing sickle cell anemia, but they did not know which exact variant drives its production.

The researchers turned to CRISPR—a gene editing tool that uses

"molecular scissors to cut DNA," according to Sanjana—to edit the regions identified by GWAS. When CRISPR edits were made at a specific location in the noncoding genome near a gene called BCL11A, it resulted high levels of fetal hemoglobin.

CRISPR has now been used in clinical trials to edit this region in bone marrow cells of dozens of patients with sickle cell anemia. After the modified cells are infused back into patients, they begin producing fetal hemoglobin, which displaces the mutated adult form of hemoglobin, effectively curing them of sickle-cell disease.

"This success story in treating sickle cell disease is a result of combining insights from GWAS with gene editing," said Sanjana. "But it took years of research on only one disease. How do we scale this up to better identify causal variants and target genes from GWAS?"

GWAS meets CRISPR and single-cell sequencing

The research team created a workflow called STING-seq—Systematic Targeting and Inhibition of Noncoding GWAS loci with single-cell sequencing. STING-seq works by taking biobank-scale GWAS and looking for likely causal variants using a combination of biochemical hallmarks and regulatory elements. The researchers then use CRISPR to target each of the regions of the genomes implicated by GWAS and conduct single-cell sequencing to evaluate gene and protein expression.

In their study, the researchers illustrated the use of STING-seq to discover target genes of noncoding variants for blood traits. Blood traits—such as the percentages of platelets, white blood cells, and red blood cells—are easy to measure in routine blood tests and have been well-studied in GWAS. As a result, the researchers were able to use GWAS representing nearly 750,000 people from diverse backgrounds to study blood traits.

Once the researchers identified 543 candidate regions of the genome that may play a role in blood traits, they used a version of CRISPR called CRISPR inhibition that can silence precise regions of the genome.

After CRISPR silencing of regions identified by GWAS, the researchers looked at the expression of nearby genes in individual cells to see if particular genes were turned on or off. If they saw a difference in gene expression between cells where variants were and were not silenced, they could link specific noncoding regions to target genes. By doing this, the researchers could pinpoint which noncoding regions are central to specific traits (and which ones are not) and often also the cellular pathways through which these noncoding regions work.

"The power of STING-seq is we could apply this approach to any disease or trait," said John Morris, a postdoctoral associate at the New York Genome Center and NYU and the first author of the study.

Using STING-seq to test clusters of likely variants and see their impact on genes eliminates the guesswork scientists previously encountered when faced with linkage among variants or genes closest to variants, which are often but not always the target gene. In the case of a blood trait called monocyte count, applying CRISPR caused one gene, CD52, to clearly stand out as significantly altered—and while CD52 was near the variant of interest, it was not the closest gene, so may have been overlooked using previous methods.

In another analysis, the researchers identified a gene called PTPRC that is associated with 10 blood traits, including those related to red and white blood cells and platelets. However, there are several GWAS-identified noncoding variants within close proximity and it was challenging to understand which (if any) could modulate PTPRC expression. Applying STING-seq enabled them to isolate which variants were causal by seeing which changed PTPRC expression.

STING-seq and beyond

While STING-seq can identify the target gene and causal variant by silencing the variants, it does not explain the direction of the effect—whether a specific noncoding variant will crank up or reduce expression of a nearby gene. The researchers took their approach a step further to create a complementary approach they call beeSTING-seq (base editing STING-seq) that uses CRISPR to precisely insert a genetic [variant](#) instead of just inhibiting that region of the [genome](#).

The researchers envision STING-seq and beeSTING-seq being used to identify causal variants for a wide range of diseases that can either be treated with gene editing—as was used in sickle cell anemia—or with drugs that target specific genes or cellular pathways.

"Now that we can connect noncoding variants to target genes, this gives us evidence that either small molecules or antibody therapies could be developed to change the expression of specific genes," said Lappalainen.

More information: John A. Morris et al, Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens, *Science* (2023). [DOI: 10.1126/science.adh7699](https://doi.org/10.1126/science.adh7699).
www.science.org/doi/10.1126/science.adh7699

Provided by New York University

Citation: CRISPR and single-cell sequencing pinpoint causal genetic variants for traits and diseases (2023, May 4) retrieved 16 July 2023 from <https://medicalxpress.com/news/2023-05-crispr-single-cell-sequencing-causal-genetic.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.