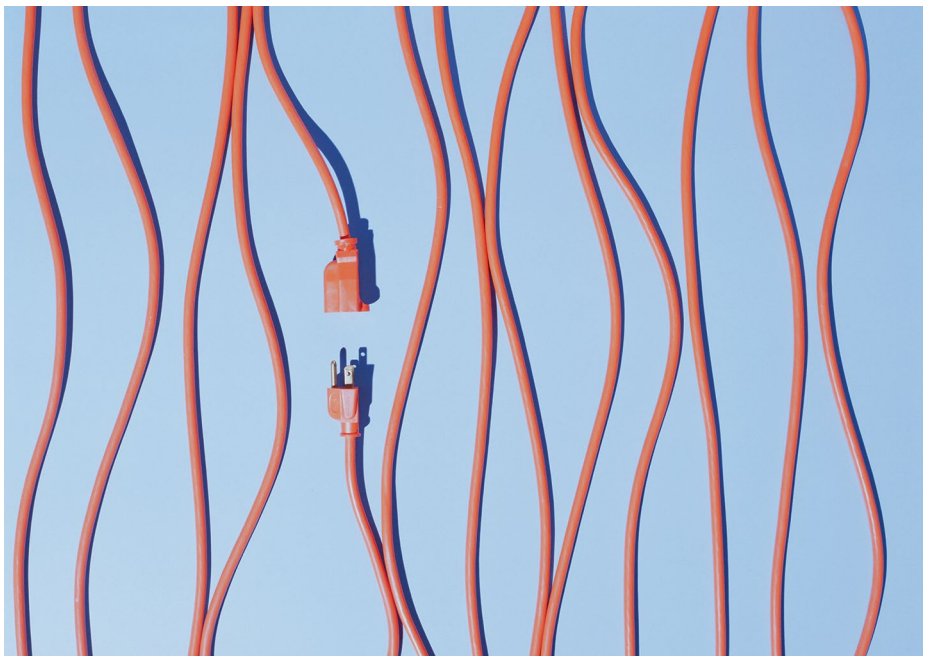
GWAS

Connecting noncoding variants to human traits



Genome-wide association studies (GWAS) have identified thousands of genomic loci associated with disease, but most reside in non-coding regions with unknown function. Now, Morris et al. present a workflow to identify target genes and pathways at non-coding GWAS loci using single-cell CRISPR screens, opening avenues for identifying drug targets for genetic diseases.

The authors developed a method termed STING-seq, which combines CRISPR inhibition (CRISPRi) and single-cell sequencing to identify GWAS variants that are causal for human traits. STING-seq was applied to study blood cell traits — including traits for platelets, red blood cells and white blood cells — in human erythroid progenitor (K562) cells, a well-characterized model for studying blood trait loci. Notably, blood traits are highly polygenic, well-studied in GWAS and have links to many common diseases.

First, the authors identified variants that might be causal from biobank-scale GWAS blood trait datasets, comprising over a million individuals of diverse ancestries. Fine mapping — a statistical method that attempts to identify plausibly causal GWAS variants — was integrated with functional datasets to further refine candidate variants. The authors ultimately prioritized 543 variants mapping to candidate *cis*-regulatory elements (cCREs).

To identify target genes for these cCREs, massively parallel CRISPRi screens were performed. The effects of CRISPRi perturbations were determined using single-cell transcriptomic and proteomic sequencing data. Negative (nontargeting) control gRNAs had no effect on gene expression or protein levels, indicating a lack of off-target effects. Of the selected cCREs, 134 target genes were identified, demonstrating the use of STING-seq for connecting noncoding variants to target genes.

Most cCREs targeted a single gene, often the nearest gene in *cis*, although a small subset of cCREs controlled several *cis* genes, reflecting either direct multi-gene regulation or indirect effects on other nearby genes driven by a single *cis*-targeted gene. Combining STING-seq with previous genome-wide Perturb-seq (GWPS) data enabled the authors to distinguish between these two scenarios, revealing, for example, that a single cCRE that was fine-mapped to be causal for an immature red blood cell trait, was likely to co-regulate *APOE* and *APOC1* genes.

A common shortcoming of GWAS interpretation is an emphasis on protein-coding genes and a dismissal of potentially relevent noncoding RNAs. A small subset of cCREs was found to control target genes in *trans*. These cCREs targeted in *cis* transcription factors (such as *GFI1B*, *NFE2*, *IKZF1*, *HHEX* and *RUNX1*) or microRNAs (such as *miR-142* and *miR-144/451*). For example, perturbation of one cCRE variant led to the detection of a 119-gene *trans*-regulatory network. Moreover, nearly all *trans*-regulatory networks studied had an enrichment for blood cell GWAS genes. These findings show that GWAS variants that map to cCREs can have effects on entire *trans*-regulatory networks.

Although STING-seq can link causal variants to target genes via CRISPRi, the method does not provide details about the direction or magnitude of effect. To gain further insight, the authors expanded their method to perform precise insertion of GWAS variants using base editing. This was achieved by fusing a cytosine base editor to a PAM-flexible Cas9 variant and performing single-cell pooled base editing (beeSTING-seq). By targeting 46 cytosine-to-thymine (C > T) GWAS variants, the authors showed that STING-seq can be used to identify the direction of effect of GWAS variants on gene expression.

> "an approach for the characterization of functional effects of GWAS loci in *cis* and *trans*."

Overall, the authors provide an approach for the characterization of functional effects of GWAS loci in *cis* and *trans*. Extending this approach to study other diseases or traits represents an important next step for target gene identification for diverse GWAS traits.

Michael Attwaters

Original article: Morris, J. A. et al. Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens. *Science* **380**, eadh7699 (2023)

CREDIT: PAUL TAYLOR/STONE/GETTY