

Scientists Used CRISPR to Trace Every Human Gene to Its Function

By **Shelly Fan** - June 20, 2022

Genes are like Egyptian hieroglyphs. Thanks to advances in whole-genome sequencing, it's increasingly easy to read each DNA letter. But the strings of A, T, C, and G bring up a second puzzle: what, if anything, do they mean?

It's a problem that has haunted biologists since the completion of the Human Genome Project. By tapping into our genetic base code, the project assumed, we'd be able to master control of inherited diseases, edit them at will, and easily predict the consequences of any gene that laid the foundation for our bodies, functions, and lives.

The vision didn't exactly work out. DNA sequences, while capturing extremely powerful genetic information, don't necessarily translate to indicating how our bodies behave. Genes can turn on or off in different tissues depending on the cell's need. Reading a DNA sequence for any gene is like parsing the base code of a cell's internal program. There's the raw genetic code—the genotype—which determines the phenotype, life's software that controls how cells behave. Linking the two has taken decades of painstaking experiments, slowly building up an encyclopedia of knowledge that decodes the influence of a gene on biological functions.

A new study ramped up the effort. Led by Drs. Thomas Norman and Jonathan Weissman at Memorial Sloan Kettering Cancer Center in New York and the University of California, San Francisco, respectively, the team built a Rosetta Stone for translating genotypes to phenotypes, with the help of CRISPR.

They went big. Changing gene expression in over 2.5 million human cells, the tech, dubbed Perturb-seq, comprehensively mapped how each genetic perturbation alters the cell. The technology centers around a sort of CRISPR on steroids. Once introduced into cells, Perturb-seq rapidly changes thousands of genes—a brutal shakeup at the genomic scale to see how single cells respond.

In other words, Perturb-seq is a large-scale tool that can help scientists translate DNA code to function—a Rosetta Stone for uncovering our cells' inner workings. Years in the making, the dataset is open for anyone to explore.

“I think this dataset is going to enable all sorts of analyses that we haven't even thought up yet by people who come from other parts of biology, and suddenly they just have this available to draw on,” said Norman.

Lost in Translation

What's the function of a gene? It's easy to think that genes are your destiny but that's far from the truth. Environmental factors, such as a massive bowl of spaghetti or a walk along the beach, can easily change gene expression, bodily functions, and potentially your body and mind.

If that's the case, what's the point of sequencing whole genomes if the outcome is always in flux? “A central goal of genetics is to define the relationships between genotype and phenotype,” the authors said. In other words, what does any gene *actually do*?

Scientists have long sought to build a bridge between genotype and phenotype. It's a painstaking process. One method, for example, perturbs genes that may be related to a disorder one by one and observes the cells' behavior. Dubbed “forward genetics,” the idea is gene-focused rather than focusing on the phenotype. An alternative approach, “reverse genetics,” dives deep into how a body or mind changes with a specific genetic edit.

Each method is an uphill struggle. With **over 20,000 genes** in our bodies and every cell behaving slightly differently (even with the same genetic changes), deciphering a gene's function often takes years, if not decades.

Is there any way to speed the process up?

The CRISPR Rosetta Stone

Enter CRISPR. Long revered as a genetic editing multitool, the method has further blossomed into a biological translator. At its heart is a technology dubbed Perturb-seq, **first published in 2016** to dissect the expression of genes. Perturb-seq makes it possible to follow the consequences of turning a gene on or off in a single cell. The method rapidly **rose to fame** in 2020 for its efficiency at altering multiple genes at once.

It's a huge win for cell biology, said the team. While scientists have readily chipped away at the massive web connecting genes and proteins, nailing down the role of individual genes has been a struggle. "We often take all the cells where 'gene X' is knocked down and average them together to look at how they changed," **said** Weissman. "But sometimes when you knock down a gene, different cells that are losing that same gene behave differently, and that behavior may be missed by the average."

The idea behind Perturb-seq is pretty simple. Imagine a toddler breaking stuff and realizing what he's done after seeing the consequences. Perturb-seq uses CRISPR-Cas9 to silence multiple genes at once, which may sometimes change a cell's behavior. While powerful, the tool has been hard to scale, studying at most a few hundred genetic perturbations at once for pre-defined biological questions.

So why not expand the method to the whole genome?

"The advantage of Perturb-seq is it lets you get a big dataset in an unbiased way," **said** Norman. "No one knows entirely what the limits are of what you can get out of that kind of dataset. Now, the question is, what do you actually do with it?"

A Cell's Life

In the new study, the team first found the magic sauce for making genome-wide changes in human cells with CRISPR. A major point was to optimize a library of guide RNAs (sgRNAs), the “bloodhounds” that track down a gene. Next, they captured cells infected with CRISPR and analyzed their gene expression. Overall, the team focused on nearly 2,000 genes. Cross-referencing changed genes with each cell’s phenotype, they then clustered genes into networks that linked to a cellular outcome.

One enigmatic gene stood out: *C7orf26*. Nixing it with CRISPR changed how a cell builds a huge molecular complex, dubbed the Integrator, which helps make molecules that control gene activity. Before Perturb-seq, *C7orf26* had never been associated with the complex before.

In another analysis, the team found a subset of genes that changes how “daughter cells” inherit the parent genome. For example, removing some genes altered the distribution of chromosomes as a cell divides. Adding or removing a chromosome can fundamentally change our biology, such as by leading to Down Syndrome.

To Norman, this aspect is the most interesting part of Perturb-seq. “It captures a phenotype that you can only get using a single-cell readout. You can’t go after it any other way.”

This database is just the start. The team is looking to use Perturb-seq on other human cell types, and all the data is available for collaboration. With the rise of Ultima Genomics, an ultra-low-cost genomic sequencing solution, single-cell CRISPR screens are likely to play an even bigger role in biotechnology, such as in analyzing the genomes of iPSCs (induced pluripotent stem cells).

To Weissman, it may even spark a shift in how we approach cellular mysteries. “Rather than defining ahead of time what biology you’re going to be looking at, you have this map of the genotype-phenotype relationships, and you can go in and screen the database without having to do any experiments,” he said.

Image Credit: Jen Cook/Chrysos Whitehead Institute

SHELLY FAN

Shelly Xuelai Fan is a neuroscientist-turned-science writer. She completed her PhD in neuroscience at the University of British Columbia, where she developed novel treatments for neurodegeneration. While studying biological brains, she became fascinated with AI and all things biotech. Following graduation, she moved to UCSF to study blood-based factors that rejuvenate aged brains. She is the ...

[Learn More](#)

FOLLOW SHELLY:

Singularity University, Singularity Hub, Singularity Summit, SU Labs, Singularity Labs, Exponential Medicine, Exponential Finance and all associated logos and design elements are trademarks and/or service marks of Singularity Education Group.

© 2022 Singularity Education Group. All Rights Reserved.

Singularity University is not a degree granting institution.