CRISPR: A Text Editor for the Human Genome

Søren Hough
June 3, 2016

Your Operating System Runs DNA

Take a breath — now let it out. What just happened? A mixture of gases from the environment entered your mouth or your nose and passed into your lungs. Oxygen from that mixture was filtered into your bloodstream, allowing your cells to make the energy you need to live. It is amazing to think that such a complex process occurs in our bodies every second of every day without any conscious input.

Perhaps more impressive than that is the idea that every aspect of breathing is guided by a fundamental code called DNA. The oxygen you're taking in? Produced by plants and bacteria through a process regulated by DNA. The food you eat every day? More often than not, it comes from plants, animals and bacteria which rely on impossibly complex DNA repositories that are being forked, edited, and modified all around us. Four chemicals known as nucleotides — denoted A, T, G, C — are the 0s and 1s that lay the groundwork for all living things on Earth.

DNA is implicated in the dark side of our lives, as well. From cancer, ALS and diabetes to obesity, depression and suicidality — the fundamental code that underpins our existence also threatens our health. Sometimes we're born with these problems. Consider non-

functioning *BRCA1*, a famous gene associated with breast cancer, which can be passed from parent to child along with a higher lifelong risk of acquiring the disease. In other cases, exposure to certain carcinogens such as tobacco smoke or excessive sunlight can lead to permanent DNA-level alterations, eventually causing cancer.

Wouldn't it be great if we could change that? What if, like a text document or a line of code, we weren't stuck with what nature gave humans? This idea has driven modern-day genetics for decades. Only recently has this become a possibility.

The Search for DNA Editing Tools

As soon as it was understood that a single code directed the formation of all living things as we know them, the race to understand and work with DNA in the context of human health was on. In 1978, restriction endonucleases — proteins which can target and cut up small predetermined sequences of DNA—helped revolutionize laboratory research and drug development. Later on, tools like zinc-finger nucleases (ZFNs) and TALENs emerged as powerful — if cumbersome — methods for targeting the genome at much larger predetermined sequences.

Fast forward to 2012, when Drs. Jennifer Doudna, Emanuelle Charpentier and Feng Zhang helped shift the field further still with the discovery and application of CRISPR. CRISPR, like ZFNs and TALENs before it, offers a new way to manipulate the genome. However unlike its predecessors, CRISPR's ease-of-use and comparatively simple development cycle means thousands of tens of thousands of researchers around the world could quickly begin to study and use the system to evaluate key genetic pathways.

Extracted from a bacterial immune system, the CRISPR system is composed of two key features: Cas9, a "molecular scissors" protein which can bind and cut DNA, and a guide (called an "sgRNA") which tells the Cas9 scissors where to go by way of a matching nucleotide sequence. The overarching principle of modern genetics is that DNA (the "code") produces RNA (the "message") and eventually protein (the "product").

Therefore, if we can delete or change the code in the cell, we can also affect the message and, eventually, its product. As it happens, scientists typically insert DNA that codes for the CRISPR features directly into the cell. The cell effectively "boots up" the Cas9 and sgRNA products from this DNA, which go on to edit the cell's own source code.

What is Possible with CRISPR?

It is important to note that CRISPR itself does not fix genes in the cell; it's only part of the editing process. Like a pair of scissors, CRISPR activity begins and ends with the nuclease's ability to cut. Sometimes this is enough. For some disease genes, their time has passed; perhaps they were essential during development in the womb, but now they do little more than generate toxic proteins. Simply cutting them out, therefore, can alleviate toxicity without impacting the organism's continued well-being.

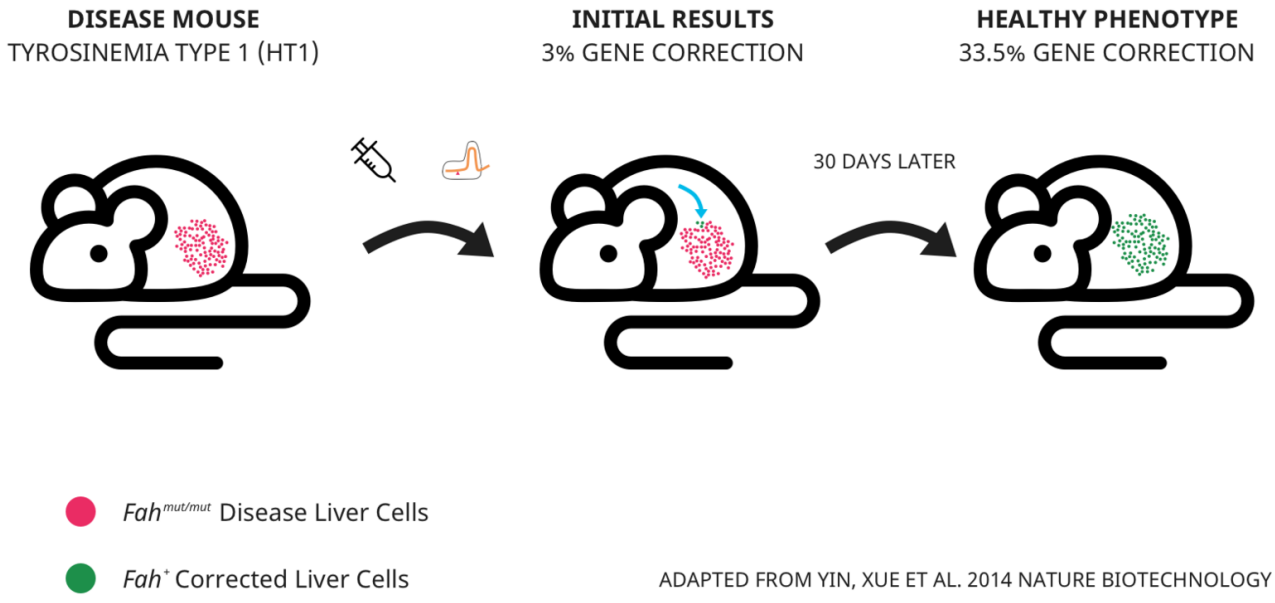
In other cases, scientists seek to fix whatever genetic mutation lead to disease. For example, *FAH* is an important gene in the liver, which, if a single nucleotide is out of place,

causes massive damage and a disease called tyrosinemia. *FAH* is useful at all phases of life, so cutting the gene out isn't an option. But what if we could swap out erroneous nucleotides in patients with that genetic disposition?

Although CRISPR itself can't repair genes, our cells have a knack for repairing any mistakes or errors in our genome. So, once Cas9 has cut a target gene, our cell immediately goes to work trying to patch up the "hole" it leaves behind. This can happen in two ways. The first is through a process called "nonhomologous end joining," or NHEJ. In this case, the cell will insert a random sequence into the space as it seeks to reconnect the DNA strands.

That doesn't work in the case of *FAH*; that gene needs to be fixed, not replaced with random nucleotides. Ideally, researchers would change the erroneous nucleotide (A) to the proper one (G) and bring *FAH* back to full working order. This is where the second pathway comes in: homology-directed repair (HDR). In HDR, the cell uses a template strand of DNA to repair the gap instead of inserting an arbitrary sequence.

This is what scientists — like those at the Jacks laboratory at MIT — have already done. By providing a corrected template for the cell to work with, they were able to employ HDR to restore the *Fah* gene to its working form in mice. This led not only to a change on the genetic level, but also a reverse in weight loss and other important indicators of tyrosinemia. This discovery represents an important step forward in therapeutic development for liver disease and, potentially, other genetic illnesses.



Concerns Around Genome Editing

With CRISPR in play, many questions continue to arise. The ethical considerations of genome editing have been and continue to be a subject of debate. The National Academy of Sciences hosted a meeting in December 2015 for the world's leading experts to discuss the ramifications of these kinds of therapies. Nevertheless, as we have discussed previously, countries take different approaches to regulating genetic technologies. While

the UK is perhaps leading the way when it comes to evidence-based policymaking around genetic technologies, the science still has a ways to go before CRISPR is ready for human trials.

One of the primary issues is that Cas9 sometimes cuts an "off-target" site when its guide inappropriately recognizes a similar sequence. Such off-target hits could render accidental gene edits, or worse, create a mutation that alters the behavior of the cell or makes it tumorigenic. Compounding this problem is the fact that the majority of researchers base the computational design of their CRISPR systems around "reference genomes" — they use standardized data models of a cell that are based on an "average" genome, as opposed to the actual genotype of the cell. In practice, every organism has a unique genome, meaning that unless we have their full genomic sequence, we may design sgRNAs that have more or different off-target sites than we initially predict.

The Importance of Multidisciplinary Collaboration

CRISPR, as with many of the biggest moments in science history, was established through interdisciplinary partnership: microbiologists extracted a useful system from bacteria and biochemists took it in a whole new direction. This is a phenomenon we've seen time and time again. Think of the discovery of DNA, which combined Watson and Crick's biologist backgrounds with Rosalind Franklin's expertise in X-ray crystallography. In molecular biophysics, it was the marriage of high-powered optics with the desire to look at individual biological molecules that led to the development of super-resolution microscopes.

As CRISPR becomes ubiquitous, the new question is how scientists can most efficiently approach the gargantuan problem of systematically discovering and evaluating the mountain of genes present in every organism. While there is a lot of hype around gene editing, designer babies and gene drives, the real value of CRISPR is the simple fact that we can turn genes on and off, one at a time or in a massively parallel fashion. This is already giving us tremendous insight into how our genomes actually work.

We know that this flow is bottlenecked by a lack of computational tools for designing and running CRISPR experiments, which is why we're doing what we're doing at Desktop Genetics. Once again, the call for interdisciplinary expertise is essential as the world's best minds work toward a better, healthier future.

Beyond this company, our entire field needs more coders and hackers — biologists and computer scientists alike — to get together and catapult biology and medicine into the 21st century and beyond. Where before we could only read our genes, now we can read and write. This is the future imagined by geneticists since the discovery of DNA, and it's going to take all of us working together to explore this brand new frontier.

Sources

Jinek M, Chylinski K, Fonfara I et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012 Aug 17;337(6096):816-21. doi: 10.1126/science.1225829. Epub 2012 Jun 28. PubMed PMID: 22745249.

Yin H, Xue W, Chen S, Bogorad RL et al. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat Biotechnol. 2014 Jun;32(6):551-3. doi: 10.1038/nbt.2884. Epub 2014 Mar 30. Erratum in: Nat Biotechnol. 2014 Sep;32(9):952. PubMed PMID: 24681508.