CREDIT: ILLUSTRATION BY MICHAEL GLENWOOD

The Hunt for Missing Genes

Identifying healthy human "knockouts"—people completely lacking a specific gene—may suggest new biomedical treatments

Daniel MacArthur's quest for the genes we can live without began with two sick boys. In 2000, as an undergraduate student, he began working in the laboratory of Kathryn North, a geneticist who studies rare, mysterious muscle diseases. Her group at the University of Sydney in Australia had recently published the possible cause of two brothers' early-onset form of muscular dystrophy. They suspected a pair of faulty copies of *ACTN3*, a gene that codes for a protein in the fast-twitch muscles that generate short bursts of power.

But later, the group tested the parents to confirm what seemed obvious: that each had two versions of *ACTN3*—a working copy and a broken one—and both had passed on the latter to their ill children. To the team's surprise, the parents, like their sons, both lacked any functioning copy of the gene or any trace of its protein in their muscles. Yet both mother and father appeared healthy.

North's group eventually did find another mutated gene responsible for the boys' disease, but at the time they also realized they had documented something significant. "It was actually incredibly exciting," says MacArthur, who joined the lab soon after the misstep was recognized. "It was one of the first examples of a gene that should be important" but that people can live without.

Nor was that family a rarity: North's group found that about 16% of the global population had two broken copies of the gene, without obvious disease as a result. As he continued in North's lab into graduate school, MacArthur and co-workers also discovered that the functional version of the gene was more prevalent in Australia's elite sprinters and jumpers.

But MacArthur became most intrigued by the prospect that hidden in the human population were people who lacked certain genes yet remained healthy. Researchers routinely disable, or "knock out," a specific gene in mice to learn what the gene does, but the results don't always translate into people. Ethically, knocking out genes in humans is off-limits. But the sick children's parents offered another route—finding natural human knockouts and looking for differences between their physiology and that of people with the intact gene. "I became fascinated by the idea that these individuals serve as experiments of nature," says MacArthur, now at Massachusetts General Hospital (MGH) in Boston and the Broad Institute in Cambridge, Massachusetts.

Those natural experiments could have biomedical payoffs. Researchers often seek new drug targets by identifying genes that cause disease when mutated and looking for molecules that can compensate. Healthy knockouts suggest a different approach: hunting for genes that, when missing, actually confer a health benefit, then trying to mimic

Published by AAAS

Select Human Knockouts

Gene	Missing protein's role	Effect of loss	Knockout frequency	Biomedical promise
PCSK9	Enzyme triggers LDL receptor degradation	Lower LDL cholesterol, reduced heart disease	Extremely rare (2% of African-Americans lack one copy)	PCSK9 inhibitors in trials for heart disease
CCR5	Cell surface receptor	Protected against HIV	1% of northern Europeans	Drug, modified stem cell transplant
ACTN3	Acts in fast-twitch muscles	Associated with reduced sprinting ability, greater endurance	16% of global population	Unknown
CASP12	Immune response to bacteria	Resistance to sepsis	Most non-Africans	Unknown
SCN9A	Sodium channel in nerve cells	Insensitive to pain	Extremely rare	New class of pain drugs

that effect by blocking the normal gene's protein. AIDS researchers have already found that certain people lacking a working gene for a specific cell surface protein suffer no ill effects and are resistant to HIV; that protein is now a drug target and figures into other anti-HIV strategies. More recently, pharmaceutical firms began developing a potential new blockbuster class of cholesterol-lowering drugs, inspired by a woman missing a cholesterol-regulating gene.

In a survey of scores of human genomes 2 years ago, MacArthur and co-workers caught a glimpse of a bigger universe of missing genes. The survey showed that the average healthy person has about 20 genes knocked out. Now, he and several other groups want to similarly comb through many more thousands of people's genomes for missing genes to seed what MacArthur calls the Human Knockout Project.

Such an effort may be the only way to fully understand the function of many of our genes, he and others contend. But identifying missing genes is only the start of the challenge. Then, investigators have to link a dispensable gene to the human knockout's phenotype—their health measures and other traits—and that information is costly and time-consuming to collect. "Getting good phenotypes makes genome sequencing look cheap (and easy)," says cardiovascular disease researcher Helen Hobbs of the University of Texas (UT) Southwestern Medical Center in Dallas.

The poster gene

Hobbs speaks from experience. About a decade ago, she and her UT Southwestern colleague Jonathan Cohen wondered if mutations that hampered or disabled a gene called *PCSK9* explained why some people in the Dallas Heart Study had unusually low levels of harmful cholesterol. Recent studies had suggested that the gene's product, an enzyme, might regulate the body's blood levels of low-density lipoprotein (LDL) cholesterol, the type that raises the risk of heart disease.

The pair's hypothesis proved to be correct. They even found one 34-year-old woman who was a *PCSK9* knockout—she completely lacked the enzyme due to mutations in both the genes for it. She had the lowest LDL cholesterol levels of all, yet was perfectly healthy. Based on the team's work, drug companies realized that blocking the *PCSK9* enzyme might lower cholesterol levels alone or in combination with the immensely popular statins.

That hunch is paying off in clinical trials of *PCSK9*-inhibiting drugs; early results suggest they can lower blood cholesterol levels by up to 57%. *PCSK9* has become the poster child for the idea that genes whose absence confers some benefit can be an extremely attractive drug target. The very existence of a healthy knockout shows that the gene isn't essential, so blocking its protein to mimic the beneficial effect should not cause harmful side effects. The human knockout is "a shortcut. You know you could inhibit" the gene's protein, says human geneticist David Altshuler of the Broad Institute.

Until recently, however, no one knew how many genes like *PCSK9* existed. That's why MacArthur did his initial survey. After moving from Australia to a postdoc at the Sanger Institute in the United Kingdom, MacArthur and his team scanned the genomes of 185 individuals who were part of a study of human genetic variation. They homed in on DNA errors that incapacitate a gene, known as loss-of-function mutations. Because not all DNA alterations do so—and many apparent DNA errors are actually sequencing mistakes—identifying true gene knockouts was a huge analytical task.

This laborious effort paid off, revealing that the average person carries about 100 incapacitated genes—and in 20 of those cases, both the maternal and paternal copies of a gene are missing, creating a complete knockout, the team reported 2 years ago in *Science* (17 February 2012, p. 823). "More than a few of us were surprised" by such a large number, says human geneticist David van Heel of Queen Mary, University of London.

Many of the more common missing genes were involved in smell; they may have been important for helping our ancestors find food, but are unlikely to affect a modern human's fitness. Others belonged to families of related genes that serve similar functions, suggesting the missing genes were not needed because the cell has backups.

Still, that left a small, but not insignificant, number of missing genes that just might protect against disease. Other scientists took notice. National Institutes of Health (NIH) Director Francis Collins remarked at a meeting that MacArthur's study suggested "a systematic, comprehensive way of identifying where the other several dozen *PCSK9*s might be out there."

Hunting for gold

Even before MacArthur's paper, other disease researchers had followed Hobbs and Cohen's example and begun looking at the genomes of people who seem protected against disease by missing genes. For example, earlier this year, Altshuler and others reported finding a gene that, when one copy

Published by AAAS

is nonfunctional, lowers a person's risk of type 2 diabetes—by a stunning 65%.

Now, MacArthur and other knockout seekers plan to widen the search. They will sequence the DNA of a large number of healthy people, see who lacks potentially interesting genes, then study whether those individuals are somehow protected—if they're less prone to heart attacks or high blood pressure, for example—or if they are unusual in some other unexpected way. This strategy "allows you to discover things you didn't know were there," says a fan of the idea, human geneticist Leslie Biesecker of NIH.

The most efficient way to do this is not to look in the general population; novel knocked-out genes will be too rare. A faster way may be to study historically isolated populations, such as the Finns. Some 4000 years ago, presumably when a small number of settlers moved to Finland, that population passed through a bottleneck. Because of this initial small gene pool, the frequency of some loss-of-function mutations is "enriched," says Aarno Palotie of the Broad Institute and the University of Helsinki. As part of a project pooling samples and data on 200,000 Finns from various biobanks, he and Mark Daly of MGH and the Broad Institute led a pilot study in which they scoured the protein-coding portions of the genome, or exomes, of 3000 Finns; they found twice as many knocked-out genes as in a northern European comparison group.

Marrying that work with DNA data and health records for 35,000 Finns has already led to potential gold: 227 individuals entirely lacked a gene called *LPA* that codes for a blood lipoprotein implicated in cardiovascular disease. These people had a significantly lower risk of heart attacks and stroke, according to a retrospective analysis that the team presented last fall at the annual meeting of the American Society of Human Genetics.

"This is a proof of concept that, indeed, you can find protective loss-of-function variants using isolated populations," Palotie says. "The excitement is that we can find many more of these." Finland's vast electronic system of medical and death records should give the team an advantage in the hunt.

Cultural traditions may also help the search for knockouts. In some cultures, it is common practice for relatives as close as first cousins to marry. That increases the odds a child will inherit two copies of the exact same nonfunctioning genetic variant, says Yale University human geneticist Richard Lifton. This has a downside if the mutations are in necessary genes—countries such as Turkey and Saudi Arabia have relatively high rates of inherited genetic diseases—but it ups the odds of finding beneficial cases.

To take advantage of this, van Heel and Richard Trembath, his colleague at Queen Mary, University of London, along with MacArthur and others, plan to sequence the exomes of up to 25,000 adults of Pakistani and Bangladeshi descent living in East London. They are now engaging with community leaders to build support for the Wellcome Trust–funded project and plan to begin the search next year. In a pilot study looking at the genomes of 1103 healthy British Pakistani individuals, they've already found about 200 potentially interesting knocked-out genes that differ from the ones found in other studies. "Some 20 are reportedly lethal in mice [if knocked out], but clearly these fit adults are doing just fine," van Heel says.

In another effort, starting this summer, a team in Saudi Arabia will collect blood samples and basic clinical data from volunteers recruited in public areas such as malls and parks—"places where healthy people are likely to go," says study leader Fowzan Alkuraya of King Faisal Specialist Hospital and Research Center in Riyadh. The goal is to enroll 10,000 people whose parents are first cousins and sequence their exomes over the next year; it is a subproject of the country's recently announced 100,000 genomes project. Like other knockout hunters, the team plans to later recall individuals with interesting loss-of-function genes for more clinical testing. "I'm pretty sure we're going to find something as exciting as the *PCSK9* story," Alkuraya says.

> "I became fascinated by the idea that these individuals serve as experiments of nature."
>
> **—DANIEL MACARTHUR,** MGH AND THE BROAD INSTITUTE

A lofty goal

Not everybody thinks that the search for knockouts should focus exclusively on healthy people. Lifton suggests that more drug targets will come from people in whom a missing gene causes a disease or some kind of obvious abnormality. He cites a gene involved in making a neurotransmittor that causes narcolepsy when both copies are disabled. Studying this gene has led to a new kind of sleeping pill, now awaiting regulatory approval. Another example is the rare individuals who feel no pain because they lack a gene for a particular cell receptor. The gene could eventually result in a new class of painkillers.

Hobbs and Cohen add that sequencing large groups of healthy people without having good clinical information to guide the search may lead nowhere. They credit their success at identifying the *PCSK9* knockout and several others to having years of detailed clinical data on people they have studied. When they find a mutation, they can immediately know a lot about the medical consequences, the pair says.

That is especially important because the consequences of missing a gene won't always be straightforward. The missing gene may have no effect in some people; or it may matter only later in life, or when the person eats certain plants or is exposed to a specific disease, Hobbs says.

MacArthur has heard the warnings, and he plans to bring together both healthy knockouts and those linked to disease, along with as much clinical data as he can amass, in one loss-of-function mutation database—the foundation, he hopes, for an eventual Human Knockout Project. Besides gathering published data on diseases caused by nonfunctioning genes, his group is building a list of novel knocked-out genes by combing through the exomes of more than 80,000 ill and healthy people sequenced for disease studies at the Broad Institute and other research centers. Eventually, he hopes other researchers, including those working in the United Kingdom, Finland, and Saudi Arabia, will contribute their knockouts, along with data on clinical consequences.

The proposal is "massively ambitious," MacArthur acknowledges, and will require buy-in from the broader genetics community. For now, he adds, "This is the beginning of an idea."

For all the voices of caution, few doubt that tallying up the world's human knockouts is a worthy goal. "We think the key is to have an outstanding, centralized database," Hobbs and Cohen note in an e-mail, "that contains systematic phenotypic information in each human knockout." Adds Lifton: "It's an obvious thing to do." **–JOCELYN KAISER**

Published by AAAS