

*The Atlantic*

# Searching for the Genes That Are Unique to Humans

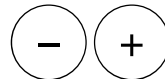
The human genome is littered with unique genes that may have been important to our evolution. And they're a bit like Oreos.



mihoda / Flickr

ED YONG | OCT 13, 2015 | SCIENCE

TEXT SIZE



The human genome contains between 20,000 and 25,000 genes. Most of these pre-date our species by millions of years, and have counterparts in chimps, mice, flies, yeast, and even bacteria. But some of our genes are ours alone. They are human-specific innovations that arose within the last few million years.

These genes might have contributed to the distinctive traits that make us human, but ironically, they are also very hard to study and often ignored. Many are missing from the reference human genome, which was supposedly “completed” in 2003.

One such unique human gene is [HYDIN2](#). It first appeared around 3.1 million years ago, as a [duplicate of an existing gene](#) called HYDIN. During the duplication process, “the head got chopped off and the tail got chopped off,” explains [Max Dougherty](#) from the University of Washington. It was as if someone had transcribed a book but neglected the prologue and epilogue. That should have been a fatal mistake since the prologues of genes contain sequences called promoters, which switch them on or off. The new gene should have been dead on arrival—a book that couldn’t be opened.

Instead, as luck would have it, it fused with a copy of *another* gene, which gave it a new lease on life. The fusion, which Dougherty described at the American Society of Human Genetics 2015 conference, created an entirely original gene, which looks like HYDIN but with a new prologue and a new first chapter. And while HYDIN, like most of our genes, exists in many other animals, its wayward daughter—HYDIN2—is a human-only innovation.

It’s also a human-*wide* innovation. Dougherty noted that despite HYDIN2’s relatively recent origins, it is now found in every living person. That’s especially surprising because it landed in a particularly turbulent part of the genome, which often gets rearranged or outright deleted. It had every chance to be lost, and yet has stood the test of time.

The original HYDIN gene plays roles in many parts of the body, including the brain. “We think that original function has been partitioned,” says [Evan Eichler](#), who led the study. He speculates that the ancestral gene still carries out its usual jobs in other tissues, which is why mutations in HYDIN lead to [a rare disease of the lungs and airways](#). Meanwhile, HYDIN2 may have taken over the brain jobs, which is why it is exceptionally active in neurons. And its origins at the very dawn

of the *Homo* lineage, before our brains ballooned to their current large size, make this potential role that much more exciting.

The team still need to confirm that the HYDIN2 gene is truly functional. But if Eichler is right, HYDIN2 would join a small but growing club of genes that arose through duplications, are unique to humans, and perform important functions in the brain. “The fact that these human-specific genes are still being discovered, years after the Human Genome Project, is pretty frickin’ amazing to me,” says Eichler.

---

**“Now we have two human-specific genes driving cognitive adaption. Pretty amazing stuff.”**

---

Eichler has been obsessed with duplicated genes ever since he was a graduate student in the 1990s. In 2002, he produced a duplication map of our DNA, [a cartography of copied genes](#). Since then, his team have been characterizing these sequences in many different species, and they’ve started to realize how weird the human ones are.

Duplicated genes make up some 5 percent of the human genome. [Many of them have arisen in the last 10 to 15 million years](#), since humans, chimps and gorillas started going our separate evolutionary ways. In fact, we—the great African apes—have ended up with far more duplicated genes than, say, orangutans or macaque monkeys. No one fully understands why.

What’s clearer is that these genes are organized in a very unusual way. For example, in other mammals like elephants, rats, and platypuses, the copies tend to sit next to the originals in a tandem series. But in humans, chimps, and

gorillas, they disperse across the genome.

They also have a unique architecture. Imagine a gene, G1, which gets copied into a different part of the genome, producing G2. Now, another duplication event copies G2, creating yet another copy of G1 *along with some of the new DNA surrounding it*. This happens again and again; with each new duplication event, the core genes picks up more flanking material. “It builds an inverse Oreo cookie,” he says, while holding his hands out and pulling them further and further apart.

The creamy filling in that Oreo is what Eichler calls the “[core duplicon](#)”—the genes that started the cascade of duplications in the first place. These tend to be rapidly evolving genes that are unique to humans and other great apes, and have probably conferred important benefits during our evolution. They are often very active, and frequently in neurons.

But “their function is a black box,” says Eichler. Unlike other older genes, you can’t tease apart their roles in simpler lab animals like mice or flies—because they don’t exist there. They can be hard to find in the first place. Most traditional sequencing methods rely on reading small sections of DNA that can then be combined into a coherent whole. But since duplicated genes are almost identical to their originals, their pieces often get mistaken for parts of their ancestors and assembled incorrectly. They are, as [as Eichler once said](#), “the geneticist’s worst nightmare.”

So, it takes a lot of work to even discover these genes, let alone divine their function. For example, in 2010, Eichler’s team identified [23 human-specific duplicated genes](#) that aren’t found in other apes. One of these, SRGAP2, has been duplicated three times, producing copies that aren’t found in the reference human genome.

The second of these, [SRGAP2C](#), is especially interesting. It emerged around 2.4 million years ago, at the time in our evolution when the human brain was

becoming distinctively bigger. And [Franck Polleux](#) from the Scripps Research Institute showed that SRGAP2C controls the growth and movement of neurons, leading to a thicker set of connections between these cells.

Marta Florio and Wieland Huttner from the Max Planck Institute of Molecular Cell Biology and Genetics [found a similar example earlier this year](#). They found that a human-specific duplicated gene called ARHGAP11b was exceptionally active in [radial glia](#), a group of stem cells that generate many of the neurons in our developing brains. [When the team activated the human gene in embryonic mice](#), the rodents developed a larger pool of radial glia, and the kinds of deep folds that are typical of a human brain.

The mice weren't smarter, but the experiment showed that ARHGAP11B could have contributed to the big evolutionary enlargement of the human brain.

"That's clear example number two for me," says Eichler. "Now we have SRGAP2 and ARHGAP11b, two human-specific genes driving cognitive adaption. Pretty amazing stuff." And if the team can prove that HYDIN2 is genuinely functional, and affects neurons, two would become three.

While these duplicated genes may have contributed to features that make us human, they might also have made us more vulnerable to diseases. When two parts of the genome are nearly identical, they can often lead to massive rearrangements, where sequences are doubled, lost, or shuffled.

This kind of genomic chaos could potentially lead to diseases, as important genes are lost or disrupted. Indeed, Eichler has shown that many duplications are associated with developmental delay, autism, schizophrenia, and epilepsy. "It's like we have these landmines that have been planted across the genome that are predisposing us to a whole cadre of neurological or developmental disorders," he says.

These genes continue to be mysterious because they sit in parts of the genome that are hard to analyze with modern techniques. So when geneticists search for variants in the genome that are associated with diseases or physical traits, they

often gloss over these duplicated genes entirely. But newer techniques are starting to solve these problems, says Eichler, including devices that can decode long stretches of DNA without having to first break them into fragments.

“The technology is almost there,” says Eichler. “I hope I live long enough to learn what these genes do.”

#### ABOUT THE AUTHOR

---



**ED YONG** is a staff writer at The Atlantic, where he covers science.

---