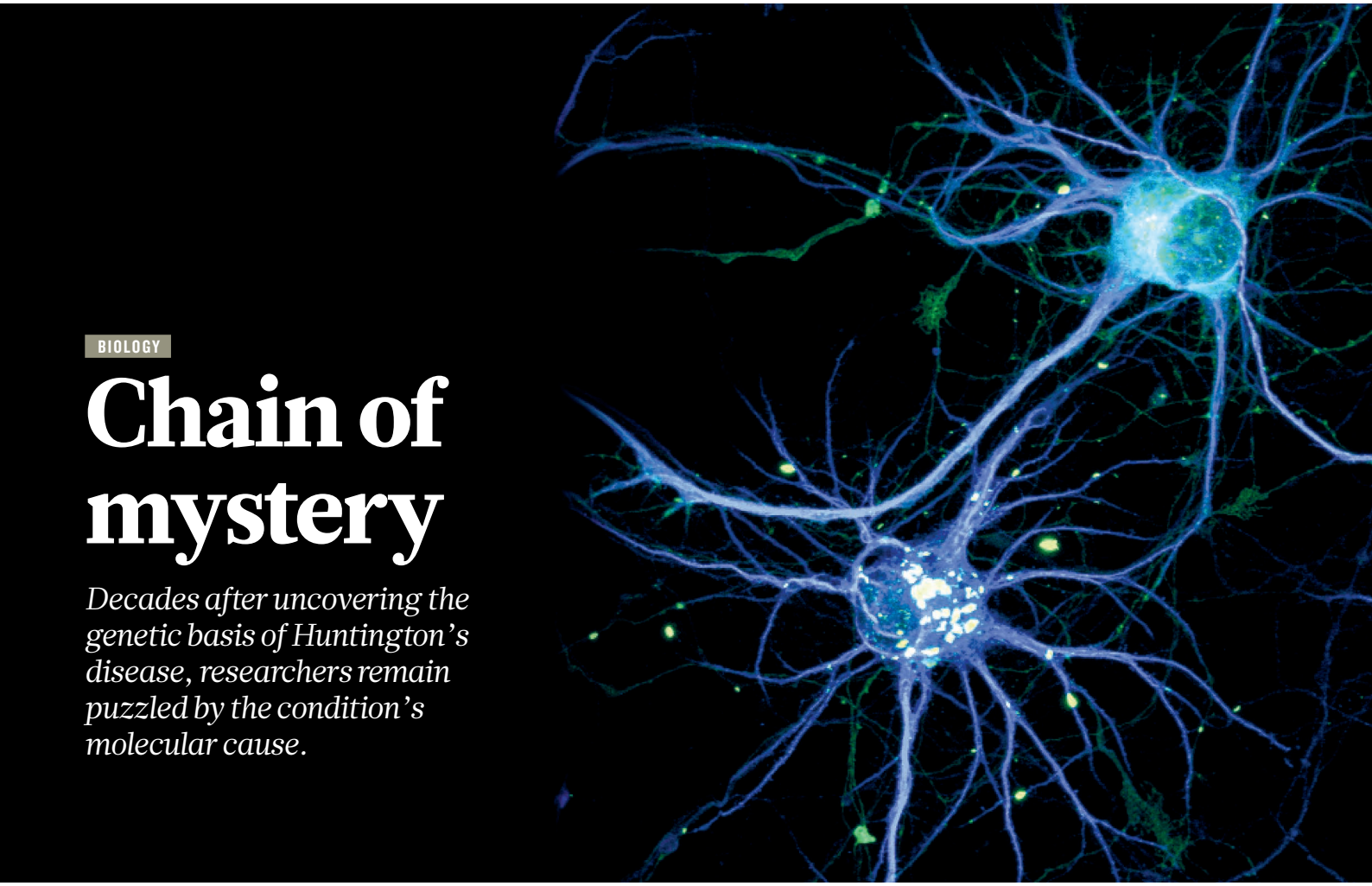


BIOLOGY

Chain of mystery

Decades after uncovering the genetic basis of Huntington's disease, researchers remain puzzled by the condition's molecular cause.

HTT exon 1 (green), a shortened form of the protein huntingtin, which is implicated in Huntington's disease, forms clumps in embryonic rat neurons.

KENNETH W. DROMBOSKY, RONALD WETZEL & TIJA C. JACOB

BY SARAH DEWEERDT

In many respects, Huntington's disease is the epitome of a well-understood hereditary disease. The condition was comprehensively described almost a century and a half ago[1]. Its symptoms are well-characterized: involuntary, jerky movements known as chorea; difficulty in coordinating voluntary movements; cognitive impairment; and psychiatric issues such as changes in mood. And its pattern of inheritance is clear — a person with a parent who is affected has a 50% chance of developing the disease.

Researchers know exactly where to find the gene that is implicated in Huntington's disease. Known as *HTT*, and located near the tip of the short arm of chromosome 4, it was the first gene to be pinned to a chromosome region through genetic mapping techniques, in 1983 (ref. 2). Ten years later, at the dawn of the genomic era, scientists homed in on the sequence of *HTT*[3]. The mutation that causes the disease is now well-understood: an abnormal expansion of a repetitive sequence of DNA comprising a triplet of bases — cytosine (C), adenine (A) and guanine (G) — towards one end of the gene. A person who carries a copy of *HTT* containing 40 or more of these triplets, or CAG repeats, will develop the characteristic symptoms of Huntington's disease, typically around the age of 45, and then succumb to the condition within about two decades of the onset of motor problems. A person who has between 36 and 39 such copies could develop Huntington's disease, but might not. Someone with 35 or fewer copies of the CAG repeat will not develop the disease.

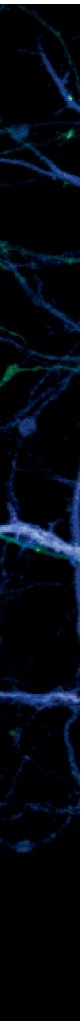
And yet, fundamental aspects of the biology that underlies Huntington's disease remain a mystery. "Many genetic diseases are relatively straightforward at the molecular level," says Ronald Wetzel, a structural biologist at the University of Pittsburgh in Pennsylvania. "The frustration we have with understanding Huntington's disease comes, in part, because of this background of our more-positive experience with other disease mechanisms."

For starters, the function of huntingtin, the protein encoded by *HTT*, isn't fully understood. More specifically, researchers don't know what the portion affected by the mutation does. And they aren't sure why the mutant protein causes problems in cells, how those problems begin, or which of several forms of the protein is responsible.

As a promising potential treatment for Huntington's disease — designed to silence the expression of *HTT* with a synthetic molecule known as an antisense oligonucleotide that binds to messenger RNA — moves into clinical trials (see page S39), answering basic questions such as these becomes more important than ever. With such progress being made, ignorance of the disease's biological basis should not be allowed to create a bottleneck for research. Knowing more about how the mutant protein leads to damage in nerve cells, in particular, could be crucial for the success of treatments for the condition. "You need to know how to most-selectively target that mRNA," says Matthew Disney, a biochemist at the Scripps Research Institute in Jupiter, Florida.

TRIPLET TROUBLE

The DNA sequence CAG encodes the amino acid glutamine. The CAG repeats in *HTT* therefore lead to the production of a string of glutamines, known as a polyglutamine chain, which is abnormally long in people with the large numbers of repeats that are associated with Huntington's disease. But the function of the chain, and the reason for the existence of



the CAG repeats, is unknown.

A somewhat provocative explanation, proposed by Chiara Zuccatto and colleagues at the University of Milan in Italy, is that Huntington's disease is the unfortunate by-product of an evolutionary advantage. The number of CAG repeats in *HTT* varies among species of vertebrate, and their expansion is greatest in humans. Huntingtin is essential for the development of the nervous system before birth; indeed, the researchers contend that CAG repeats might have contributed to the development of the complexity of the vertebrate brain[4].

But other researchers say that CAG repeats are just as likely to have arisen by chance, and that they have little effect until a threshold is crossed. "These sequences are difficult to copy," says Cynthia McMurray, a biochemist at Lawrence Berkeley National Laboratory in California, "so sometimes the polymerase [the enzyme responsible for copying DNA] stutters around, and it can't get through it." This results in extra copies being incorporated into the DNA, lengthening the repeat.

Ten hereditary diseases are known to be caused by the expansion of CAG repeats in various genes, and other triplet-repeat diseases exist — CAG repeats just happen to be particularly vulnerable to these polymerase 'stutters'.

It's not just the purpose of the polyglutamine chain that is difficult to unpick. The function of huntingtin is poorly understood and the protein might have multiple roles in cells. Although it is expressed most strongly in the brain, huntingtin is found throughout the body and has been shown to interact with more than 100 proteins. "That's sort of amazing," says Stefan Kochanek, a gene-therapy specialist at University Hospital Ulm in Germany. "It's a very large number of proteins."

Some experiments suggest that huntingtin helps to transport proteins around the cell. Other findings hint that it might play a part in transcription. It might also aid the proper folding of proteins, or help proteins to form complexes with one another. "The idea of there being one prime mechanism for what huntingtin protein does is far from clear," McMurray says. "It does a lot of things."

Scientists have documented numerous biochemical abnormalities in animal models of Huntington's disease, as well as in people with the condition. "There are probably more things you can measure that are going off the rails in Huntington's than stay on the rails," says David Housman, a biologist at the Massachusetts Institute of Technology in Cambridge, who was involved in identifying *HTT*. The balance between protein synthesis and degradation, the function of cellular structures such as the endoplasmic reticulum, which is involved in protein processing, and the cell's responses to stress, for example, become disrupted. But Housman says that the cause-and-effect relationship between these observations, or what goes awry first, is unclear. That makes it difficult to glean clues about huntingtin's normal function from observations of what goes wrong in Huntington's disease — as well as to design a treatment to target the initial steps in the development and progression of the condition.

An analysis of the structure of huntingtin[4] published in March supports the idea that it is some kind of molecular chaperone — a protein that aids the function of other proteins — says Kochanek, who led the work. It took a decade to come to fruition, he adds, because the protein's extreme flexibility made it difficult for researchers to get clear images of the structure.

The analysis does not, however, clearly reveal the structure of the polyglutamine chain, so it cannot provide further information about the function of that part of the protein. Some researchers suspect that the portion of huntingtin in which the disease-causing mutation occurs might have little involvement in the protein's normal function — a molecular curlicue, just along for the ride.

"The idea of there being one prime mechanism for what huntingtin protein does is far from clear."

But others contend that it must do something, even though it's not clear what. Polyglutamine chains often facilitate interactions between proteins, and mice that lack this portion of *HTT* show behavioural and biochemical changes. So the question remains: how exactly does mutant huntingtin cause such problems?

STICKY CHAINS AND CLUMPS

Scientists have debated whether the damage to cells in Huntington's disease stems from a loss of the normal function of huntingtin, a toxic effect that is unique to the mutant protein, or a combination of both. There's also disagreement on which form of the mutant protein is the culprit — a question that is difficult to answer because many versions of huntingtin exist in the cell at the same time, owing to splicing and modification processes. "It's like having too many viable suspects in an Agatha Christie novel," Wetzel says.

Some researchers think that full-length huntingtin is the prime suspect. Chains of glutamine tend to be sticky, McMurray says, and huntingtin containing too many glutamines in a row might adhere more strongly to other molecules than it would normally — gumming up the movement of proteins through the cell. She suggests that the resulting impairment of huntingtin's transportation activity would disrupt metabolism and other cellular functions gradually, which is consistent with the slow progression of Huntington's disease.

But many others instead point to a shortened form of huntingtin, known as HTT exon 1, in which exon 1 refers to the protein-coding region of *HTT* that contains the CAG repeats. It's found only in people with Huntington's disease and contains sequences that are not usually translated into protein. HTT exon 1 could be more toxic than any other form of huntingtin. For example, mouse models in which only HTT exon 1 is expressed show all the main features of Huntington's disease[5]. Meanwhile, studies in fruit flies suggest that of all huntingtin fragments produced by the cell, HTT exon 1 is the most harmful[6].

If HTT exon 1 is the real cause of Huntington's disease, exactly how it exerts its damaging effect remains unclear. That's because, compared with flexible full-length huntingtin, HTT exon 1 is a downright shape-shifter. It switches freely and rapidly between multiple conformations, making it almost impossible to isolate the effects of any particular one.

Researchers have several hypotheses. One idea is that, similar to how full-length huntingtin containing many glutamines might gum up cells, the longer the polyglutamine chain of HTT exon 1 is, the more strongly the fragment can bind to other molecules. Other scientists blame the tendency of HTT exon 1 to aggregate with itself. Nerve cells from people with Huntington's disease contain clumps of huntingtin-related proteins, as well as smaller fibrils comprising thousands of copies of HTT exon 1. In turn, such aggregations can disrupt a variety of cellular functions, leading inexorably to neurodegeneration.

Similar clumps are found in other neurodegenerative conditions such as Alzheimer's disease, Parkinson's disease and amyotrophic lateral sclerosis. They also occur in all known diseases caused by expanded CAG repeats, including spinocerebellar ataxia.

Wetzel and his collaborators have shown that polyglutamine chains acquire a much greater propensity to stick to each other and form aggregates when they reach 37 amino acids long. This observation provides a potential link between the disease-causing threshold of the CAG repeats and the biochemistry of huntingtin. Still, others caution that huntingtin clumps might be the result, not the trigger, of the processes responsible for the slow progression of Huntington's disease.

It's tempting to think that such debates won't matter if treatments such as the antisense oligonucleotide in clinical testing can simply turn off huntingtin expression altogether. But some of the antisense molecules being tested target a portion of *HTT* that is distant from exon 1, Disney says. "Those antisense oligonucleotides are going to ablate full-length huntingtin, but they are not going to affect this mini version of huntingtin," he says. Perhaps it's a hint that huntingtin might be too complex for such simple solutions. ■

Sarah DeWeerdt *is a freelance science writer in Seattle, Washington.*

1. Huntington, G. *Med. Surg. Rep.* **26**, 317–321 (1872).
2. Gusella, J. F. *et al. Nature* **306**, 234–238 (1983).
3. The Huntington's Disease Collaborative Research Group. *Cell* **72**, 971–983 (1993).
4. Guo, Q. *et al. Nature* **555**, 117–120 (2018).
5. Mangiarini, L. *et al. Cell* **87**, 493–506 (1996).
6. Barbaro, B. A. *et al. Hum. Mol. Genet.* **24**, 913–925 (2015).