The Puzzle of the Triple Repeats

The discovery that an unprecedented new type of mutation causes several human genetic diseases has fascinated geneticists by presenting them with a host of questions to answer

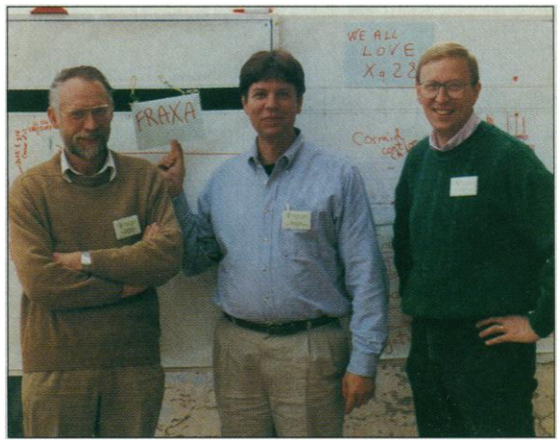
Two years ago, when researchers discovered the gene that causes a hereditary form of mental retardation known as fragile-X syndrome, they also turned up a mutation so unexpected geneticists are still scratching their heads over it. The defect, which makes genes balloon in size by adding extra copies of a three base-pair repeated sequence of DNA, was the first of its kind. Despite decades of study, nothing like it had ever been seen in any of the species that laid the foundations for modern genetics: bacteria, the fruit fly *Drosophila melanogaster*, and the mouse. "Finding a mutagenesis mechanism [in humans] that hadn't been described years ago in *Drosophila* or some other organism is unprecedented," says Stephen Warren, a human geneticist at Emory University's Howard Hughes Medical Institute who was one of the codiscoverers of the fragile-X gene.

Finding a human mutation not present in the animal models was the first surprise. The next shock followed quickly: The mutations caused by these expanding trinucleotide repeats turned out to be common causes of human disease. In the past 2 years, they have been fingered as the culprits in three hereditary disorders besides fragile-X syndrome: myotonic dystrophy, spinobulbar muscular atrophy (also known as Kennedy's disease), and—just this March—Huntington's disease. And that will not be the end of the list. The human genetics community is currently abuzz with rumors that there will soon be announcements that at least two more illnesses are caused by trinucleotide repeat expansions, although the identities of the diseases are still being closely held.

As a result, these mutations have quickly gone from being an anomaly to being one of the hottest topics in human genetics. "No one expected that DNA sequences could be so unstable or behave as these do," says Jean-Louis Mandel of INSERM in Strasbourg, France, another codiscoverer of the fragile-X gene. "And because they are exceptions—they can't be explained by classic Mendelian rules—everyone is very excited."

Since the classic Mendelian rules are the basis of our understanding of genetic inheritance, it's clear how fundamental the implications of the new findings are. In fact, much of the excitement Mandel refers to stems from the many intriguing questions the existence of the trinucleotide repeats and their mutations raises. Those questions haven't been answered yet, but when they are, the solutions could reveal essential new information about the role of genetics in human disease as well as about how the human genome operates normally.

Although the disease-causing expansions of the trinucleotide repeats have so far been found in just a handful of human genes, the trinucleotide repeats themselves are very common. At least 50 human genes are known to contain stretches of five or more triplet repeats, and researchers suspect that many others also have them. The sequences are also found in every animal studied, although they apparently don't balloon in size in other species as readily as they do in humans.



Talking it over. Ben Oostra *(left)*, Stephen Warren *(middle)*, and David Nelson discuss the intriguing new mutations.

Why these repeats are so unstable and prone to both expand and, occasionally, shrink in size in humans is one of the most baffling questions the researchers are facing. Answering it would, of course, help explain how the associated genetic diseases arise. Finding out what makes the sequences unstable might also provide a better understanding of the mechanisms that human and other cells use to maintain the stability of their genetic material as it's handed down from one generation to the next.

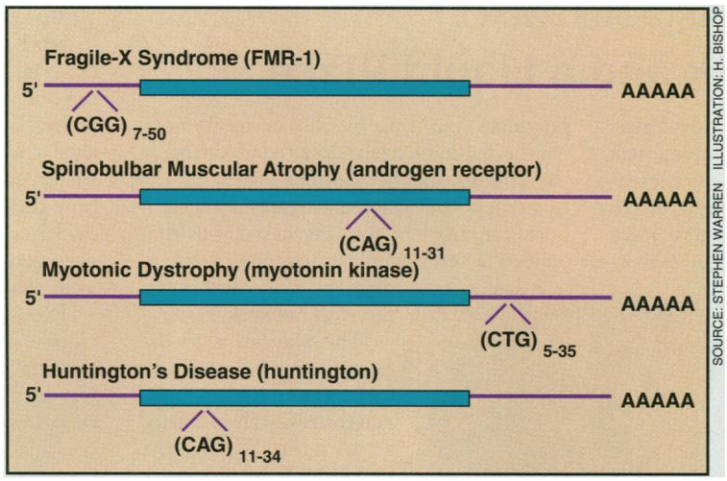
The FMR-1 gene, which is the one at fault in fragile-X syndrome, shows just how much the trinucleotide repeats can expand. The normal gene carries at most 50 copies of the CGG trinucleotide. (C stands for the base cytosine, G for guanine, A for adenine, and T for thymine.) When the gene begins to mutate that goes up to between 50 and 200 copies, although at that point the gene carriers are still healthy. But in children who inherit the gene from these carriers and actually develop mental retardation and the other fragile-X symptoms, the FMR-1 gene may have hundreds to thousands of CGG repeats. Huge expansions of another trinucleotide repeat (CTG) can also occur from one generation to the next in the gene that causes myotonic dystrophy (DM), while smaller, although no less devastating, expansions in the CAG trinucleotide repeat lead to Huntington's and Kennedy's diseases. "It's that increase from one generation to the next that is so remarkable," says Warren. "But we have yet to find a mechanism that would explain it."

The researchers have some ideas about what may be causing the problem, however. One possibility is that the expansions are linked to aberrant DNA replication either during the meiotic cell divisions that produce the egg and sperm or in the rapidly dividing cells of the early embryo. Or, suggests David Nelson of Baylor College of Medicine, the extra repeats might be inserted because of an error in DNA repair. The repair system might get confused by the repeats, he says, so it ends up getting stalled and adding too many copies. Or the replicating and repair enzymes may be overactive.

Nor do researchers know why the sex of the parent transmitting the gene affects the trinucleotide repeat expansions. In Huntington's disease, for example, the enlargement usually occurs when the gene is transmitted by the father, whereas in fragile-X, the amplification happens when the gene comes from the mother. "It may well be that this has something to do with the different biological mechanisms involved in generating sperm and eggs," says Warren. "If we could discover when the instability occurs—if it's present in the ovum prior to conception or happens after conception—we would have a better handle on many of these questions. In either case there is a great deal of instability in the somatic cells of the embryo, while adults are relatively stable. And this suggests that there is a window early in embryogenesis when the instability occurs."

However and wherever the expanded repeats form, recent evidence suggests they can also shrink back to the normal size range, although this happens rarely. Children of myotonic dystrophy patients have occasionally been found whose DM genes have shorter repeats than those of their parents.

But the overall trend is for these repeats to expand—an event that seems to be triggered when they reach a certain size threshold, says Richard H. Myers of Boston University School of Medicine. In his investigations of what appeared to be new Huntington's cases, he found that even though the patients' families had been free of the disorder for several generations, they were nevertheless passing along the chromosome that carried the gene with the unstable repeats—which eventually expanded into the disease range in some descendants. "At some point, in some generation, the gene will amplify again into the clinical range," Myers says. Such silent carriers, he suggests, may explain why these diseases, particularly those like myotonic dystrophy in which patients seldom have children, do not die out in the population.



SOURCE: STEPHEN WARREN ILLUSTRATION: H. BISHOP

Where they're located. Trinucleotide repeats occur at different sites in genes. The numbers give their normal size ranges and the solid bars indicate the genes' protein-coding segments.

The instability of the mutations also gives a rational handle for understanding a few other old problems in genetics. It's well known, for example, that hereditary diseases may display what's known as "incomplete penetrance," which means that not everyone who inherits a disease gene actually gets sick. And for those who do come down with the disease, the symptoms may vary widely from one individual to another. With the triplet genes, this variability may occur, Nelson explains, "because the mutation itself is not the same." Patients with the longer repeats would presumably have more severe symptoms than those with the shorter ones.

In addition to trying to understand the role of the trinucleotide repeats in disease, the researchers also want to know what they do normally. The fact that they have been conserved in evolution indicates that they do have a normal role, Warren says, but nobody knows what it is, especially on the global scale of the genome as a whole.

On the smaller scale of individual genes, the repeats apparently serve different purposes depending on their location. They are situated within the protein-coding portions of some genes, including the one for Huntington's, although researchers don't yet know that gene's normal function or how it's altered by the expansion of the triplet repeat. In contrast, the myotonic dystrophy gene's repeats are located not in the protein-coding segment of the DNA, but in noncoding sequence that follows it. When the repeats balloon up, the amount of protein made may be altered.

It's clear that discovery of the trinucleotide repeat mutations has created many more mysteries than it's solved. And finding the solutions is proving nearly as difficult as the now legendary searches for the fragile-X and Huntington's genes. Investigators don't have animal models for the diseases, and efforts to create systems for studying the genes by cloning them and introducing them either into cultured cells or experimental animals like mice have been hampered by the very instability of the repeats, particularly the lengthy ones in myotonic dystrophy and fragile-X. "The [cloning] hosts always lose most of the repeats, and this has hamstrung our ability to do many things we'd like," explains Nelson. He and Warren speak wistfully of the "wonderful experiments" they have developed on paper to investigate their questions. But until they discover a way to clone the repeats and create an animal model with the mutation, their experiments will stay on the shelf.

> **"No one expected that DNA sequences could be so unstable or behave as these do."**
>
> **—Jean-Louis Mandel**

Researchers working on the rare Kennedy's disease have met with greater success, perhaps because the mutant repeats number only in the 40 to 60 range, rather than in the thousands as in fragile-X. "We have successfully cloned the extended repeat, which is on the androgen receptor gene, and have also created transgenic mice," says Kenneth Fischbeck at the University of Pennsylvania. "Since we can't find the mutation in other organisms, we have to reproduce models like these—it's the best way to get at the mechanism." Even though the repeated sequences in the other genes are much longer, the mechanism underlying their mutation is presumed to be the same. Consequently, Fischbeck's transgenic mice may also supply answers to his colleagues' basic questions about where and when the expansion begins, and what triggers it. "If we can find out why one of these sequences is unstable, that may explain the others," says Warren, "and then it would no longer be necessary to clone the large repeats of fragile-X or DM."

Additional clues to the expansion mechanism may also come from the discovery of similar expanded repeat disorders, and more of these will come to light as a result of a new technique just reported in *Nature Genetics* by Martin Schalling, who did the work while in David Housman's group at the Massachusetts Institute of Technology. Until now, the mutations have been found only after molecular geneticists have identified particular disease genes and sequenced them to find out where they went wrong. But the technique, called repeat expansion detection (RED), allows researchers to search the human genome directly for large trinucleotide repeats. Mandel describes the RED technique as "elegant" and predicts it will "considerably shorten the search for such disease genes." Already with RED, Schalling's group has identified another long repeat on chromosome 18 in three families. Although they have not yet associated a disease with it, the repeat is similar to those already discovered, particularly in its tendency to grow in size from one generation to the next.

All in all, the trinucleotide repeat expansions have shaken the field of genetics almost as much as did Barbara McClintock's jumping genes, or transposons. Further, as Warren points out, "their discovery gives us one of the best arguments for the value of the Human Genome Project." When the project was first initiated, Warren and others observe, critics complained that nothing much beyond additional new genes would be found. "But this new mutation is evidence that there are probably other mechanisms of disease out there that we also haven't contemplated," says Warren. "It could be that this is just the tip of the iceberg."

–Virginia Morell