



# The director's cut: how CAG repeats change the editing of genetic messages

Scientists in Massachusetts have recently advanced our understanding of how repetitive sequences in DNA can disrupt the creation and editing of genetic messenger molecules in cells, and how this could lead to the production of harmful proteins.

By [Lucy Coupland](#) | March 26, 2024 | Edited by [Dr Rachel Harding](#)

---

**L**ong repetitive sequences of C-A-G letters in the DNA code are associated with at least 12 genetic diseases, including Huntington's disease (HD). A group of scientists in Massachusetts, USA, have recently developed a new genetic strategy to study how CAG repeats can lead to harmful proteins being made in cells, causing cells to become unhealthy. Their findings showed that expanded CAG repeats can interfere with a process called 'splicing', which chops up and organises genetic message molecules before they are turned into proteins.

## CAG repetition

Our DNA is a genetic code that holds instructions for making thousands of different proteins, the molecular machines that run our cells. This code is made of four building blocks or 'bases': C, A, G, and T. DNA is arranged like a twisted ladder with two DNA strands bound together in a helix, each made of a string of bases. The bases on one DNA strand pair with bases on the opposite DNA strand to form the 'rungs' of the ladder.



*DNA is structured like a ladder with two strands of genetic material bound together in a double helix, each made up of a sequence of letters of the genetic code. Letters on one DNA strand pair with letters on the opposite strand to form the 'rungs' of the ladder.*

HD is known as a 'CAG repeat expansion disease'. Everyone has a repetitive sequence of C-A-G DNA letters in their huntingtin gene, but people who go on to develop HD have over 36 C-A-G repeats. The number of CAG repeats can increase over time, called repeat expansion, and this seems to happen mainly in cells that get the most unhealthy in HD such as brain cells.

If we can understand exactly how a longer CAG repeat itself makes cells sick, we may be able to keep brain cells healthy and delay when HD symptoms appear. There are also other diseases caused by expansions in CAG repeats, including spinocerebellar ataxias and myotonic dystrophies. Trying to find similarities between what happens in cells affected by these other diseases may help us learn more about what goes on in HD.

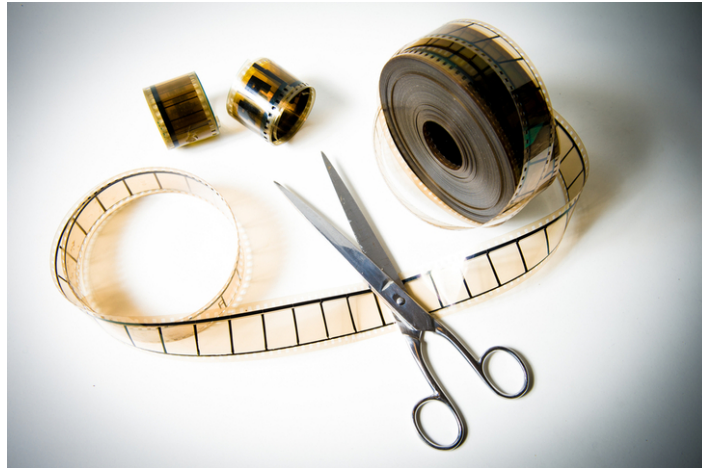
## **Cutting scenes in the genetic script**

When a cell wants to make a protein coded by a certain gene, the two DNA strands unwind and separate from each other. Cellular machinery then reads the opened-up DNA base code and makes a copy of it, called an RNA message molecule, a bit like making a photocopy of a recipe from a cookery book.

However, before any RNA message molecules are read by the next set of cellular machinery to make the corresponding protein, an essential process needs to take place. Much like editing out unnecessary scenes from a film to make a final polished version, this process involves editing the RNA message to remove all of the waffly bits of genetic code copied from DNA which aren't actually needed to make a protein. The process of going from the unedited RNA message molecule to a shorter more succinct message is called 'splicing'.

During splicing, non-essential sections of the unedited message are cut out and the important sections that remain are pasted together to produce what is known as 'mature' RNA. This final mature RNA product has only the necessary instructions that the cell needs to make proteins.

## Expanded CAG repeats can cause genetic plot twists



*RNA splicing is a crucial process in cells where certain parts of an RNA message molecule are cut out and removed, while the remaining segments are joined back together. This final RNA message has only the necessary instructions that the cell needs to make a protein. Think of it like editing a film reel, where unwanted scenes are cut out, and the remaining scenes are combined to create the final movie.*

In diseases caused by expanding CAGs, the CAG repeat in the DNA is copied into the RNA message, which can cause abnormal proteins to be made. In the case of HD, an extra-long version of the huntingtin protein is made. A group of scientists led by Dr Jain in Cambridge, Massachusetts, previously found that repeat-containing RNA messages, along with the proteins made from them, combine to form toxic clumps in cells which can cause serious damage.

To find out exactly how longer CAG repeats cause the production of harmful RNA and proteins, Rachel Anderson and colleagues within the Jain team recently developed a clever new method to look in detail at the precise genetic message in RNA molecules containing large CAG repeats. Interestingly, they found that CAG repeats in RNA cause mistakes to be made during splicing of that RNA message molecule. Expanded CAG repeats in RNA cause other sections of the message molecule, sometimes far away from the CAG repeat itself, to be cut and pasted into or next to the repeat during splicing.

Here, the expanded CAG repeat can act like the opening credits of a film, into which the final scenes of the film get mistakenly inserted out of order. When this happens, the plot of the film no longer makes sense. Similarly, the final RNA message doesn't make much sense

when other sections of genetic information are inserted into the CAG repeat during splicing. This leads to the creation of many different repeat-containing mature RNAs with unexpected sequences.

The researchers found that the longer the CAG repeat in the RNA message, the more faulty splicing events that occurred. This is interesting as the CAG number in HD tracks with the age at which symptoms start and the rate at which they progress. The researchers showed that when they stopped all splicing events in cells using a chemical, repeat-containing RNA messages did not form clumps in cells and so did not cause cell toxicity.

## Protein production glitches

So far, these results explain how expanded CAG repeats lead to abnormal and incorrectly spliced mature RNA messages, but what happens when these messages are read to make proteins? Any mature RNAs that are ready to be read by cellular machinery to make a protein contain a 'start' signal, like a green traffic light. The researchers found that sometimes when repeat-containing RNAs are incorrectly spliced, more of these start signals are found before the repeat, causing many different proteins to be made from a single RNA message than normal. The researchers altered these start signals in the CAG repeat-containing RNAs to turn them off and found that this stopped abnormal proteins from being made.



*RNA messages that are all set to be read by cellular machinery to make a protein contain 'start' signals like a green traffic light. When CAG repeat containing RNAs are edited during splicing, start signals can be incorrectly cut and pasted to the CAG repeat, causing more abnormal proteins to be made from the RNA message than normal.*

*Image credit: [Friva](#)*

The researchers also studied the RNA messages containing CAGs that were copied from genes associated with CAG repeat expansion diseases, including spinocerebellar ataxia and myotonic dystrophy. The researchers showed that expanded CAGs copied from these genes also caused abnormal splicing into the repeat, which again contained more protein reading start signals which may cause more abnormal proteins to be made.

# What does this mean for CAG repeat expansion diseases?

Understanding how important processes in cells are impacted by long CAG repeats can help researchers piece together exactly how cells become unhealthy in CAG repeat expansion diseases and point to which processes can be targeted with therapeutics. The findings from this study add another piece to the puzzle of what happens in cells, suggesting expanded CAG repeats in RNA interfere with splicing, which can lead to damaging proteins being made.

Importantly, these experiments were performed in cell types, such as kidney cells, which are easy to grow and manage in the lab but are not most affected by HD. Therefore, these cells may not accurately reflect what causes cells to become sick in HD. A lot more work is needed looking at how expanded repeats alter RNA splicing and protein production in cell and animal models of HD. Nonetheless, targeting splicing may be a potentially exciting avenue that researchers can pursue to develop medicines for HD and other repeat expansion diseases.

---

*The author and editor have no conflicts of interest to declare. [For more information about our disclosure policy see our FAQ...](#)*

---

## GLOSSARY

**spinocerebellar ataxia** A family of diseases which result in characteristic movement disorders. Many types of spinocerebellar ataxia are caused by the same type of mutation as HD – a CAG expansion.

**huntingtin protein** The protein produced by the HD gene.

**therapeutics** treatments

**CAG repeat** The stretch of DNA at the beginning of the HD gene, which contains the sequence CAG repeated many times, and is abnormally long in people who will develop HD

**splicing** the cutting up of RNA messages, to remove non-coding regions and join together coding regions.

**RNA** the chemical, similar to DNA, that makes up the 'message' molecules that cells use as working copies of genes, when manufacturing proteins.

---

© HDBuzz 2011-2024. HDBuzz content is free to share, under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

HDBuzz is not a source of medical advice. For more information visit [hdbuzz.net](https://hdbuzz.net)

Generated on May 04, 2024 — Downloaded from <https://en.hdbuzz.net/361>