

# Decoding mystery sequences involved in gene regulation

July 10 2013

---

Every cell in an organism's body has the same copy of DNA, yet different cells do different things; for example, some function as brain cells, while others form muscle tissue. How can the same DNA make different things happen? A major step forward is being announced today that has implications for our understanding of many genetically-linked diseases, such as autism.

Scientists know that much of what a gene does and produces is regulated after it is turned on. A gene first produces a molecule called RNA, to which tiny proteins called RNA binding proteins (RBPs) bind and control its fate. For instance, some of these proteins cut out parts of the RNA molecule so that it makes a particular protein, while other RBPs help destroy the RNA before it even produces a protein.

But these mechanisms are not well understood because the RNA sequences, which the RBPs bind to, have been so difficult to decipher. To fully understand [gene regulation](#) (and dysregulation, as in the case of disease), scientists have needed to employ advanced lab techniques and data analysis to identify the patterns of the RNA sequences.

This gap in knowledge motivated a team of researchers co-led by Senior Fellow Tim Hughes (University of Toronto and the Canadian Institute for Advanced Research) to produce the first-ever compendium of RNA-binding sequences, which was published in *Nature* on July 11, 2013.

"It took us a long time to generate and analyze the data," explains

Hughes. "After spending years developing and perfecting a method, we started looking at all the proteins in humans, [fruit flies](#) and other complex [organisms](#) that look like they may bind RNA and found which sequences they like to bind to. Our compendium of RNA-binding sequences will become a resource for researchers in this field, and will be especially useful in human [genetic analysis](#)."

The team found that humans and fruit flies have similar RBPs, since they derive from a [common ancestor](#), and that in many cases they essentially bind the same sequences. The researchers anticipate that this is the case for proteins in other organisms.

"We looked at just over 200 proteins in total, but can probably infer the preference for tens of thousands of proteins in many other organisms," says Hughes.

In addition, many of the sequences similar across species were at the end of the RNA transcript, which is a region associated with regulation of RNA decay or movement of the RNA to another part of the cell. "This indicates that there is probably more regulation of gene expression itself at the level of stability or destruction of RNA," explains Hughes.

One of the major insights that came out of the team's analyses was about a well-studied protein called RBFOX1, which was already known to have a function in regulating RNA splicing and to be decreased in autism. The team's findings suggest that RBFOX1 has a role in regulating the expression level of nervous-system-related genes in brains with [autism](#), and that it does so by making RNA more stable.

The underlying causes of disease are more complicated than a single gene not working right, says Hughes. He anticipates that the team's compendium will be useful in human genetic analysis.

"What often happens is that scientists identify a genetic variation associated with a disease, but then they don't understand why it leads to the disease. What exactly do these sequence changes cause? If the sequence is in a regulatory region of the RNA, then with our compendium, other scientists will be able to see what protein binds to it. This will give them a better idea of what is being disrupted."

The study was a large collaborative effort, supported in part by CIFAR, that involved Senior Fellows Brendan Frey (U of T) and Andrew Fraser (U of T) and Global Scholar Alumnus Matthew Weirauch (Cincinnati Children's Hospital Medical Center) in CIFAR's Genetic Networks program. Hamed Najafabadi (U of T), a postdoctoral fellow who performed much of the analysis in this study, was partially funded by CIFAR.

"Members of the Genetic Networks program have motivated us to look at roles for RNA-binding proteins causing dysregulation of gene expression in disease," says Hughes. "We anticipate that this new knowledge will be valuable to other program members working on specific disorders."

The next steps for the team are to expand their compendium to encompass all complex organisms.

Frey also hopes to take these findings further to build models that will more accurately describe observed gene expression patterns.

"My research focuses on deciphering the regulatory sequences in DNA, which ultimately shape the [fate](#) of an RNA molecule," explains Frey. "I hope to take the RNA-binding sequences identified in this paper and use them as tokens to figure out how they act in a regulatory fashion. This will help us better understand human disease by providing insights into how a mutation in DNA affects regulation."

**More information:** A compendium of RNA-binding motifs for decoding gene regulation, *Nature* July 11, 2013.

Provided by Canadian Institute for Advanced Research

Citation: Decoding mystery sequences involved in gene regulation (2013, July 10) retrieved 24 April 2024 from <https://phys.org/news/2013-07-decoding-mystery-sequences-involved-gene.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.