

NEWS

# New resource catalogs RNA-binding sites of many proteins

BY VIRGINIA HUGHES

12 AUGUST 2013

A new online database lists the likely RNA-binding sites of more than 8,000 proteins from 289 species, ranging from mosses to monkeys. Researchers debuted the resource in the 11 July issue of *Nature*<sup>1</sup>.

At least a thousand human proteins are thought to link up with RNA, the molecule that carries instructions to help turn DNA into protein. These so-called RNA-binding proteins, or RBPs, are powerful: They can stabilize, change or destroy RNA, controlling the way genes are expressed in a cell.

Some RBPs have been linked to autism. For example, the protein missing in **fragile X syndrome**, FMRP, **binds to the RNA of some 900 genes** and represses their expression. And some people with autism carry genetic glitches in an RBP called **RBFOX1**<sup>2</sup>.

Despite their biological importance, few RBPs have been studied in depth, however.

“Understanding how gene regulation works is probably one of the major challenges for doing anything useful with the human genome,” says **Tim Hughes**, professor of molecular genetics at the University of Toronto and lead investigator of the new study. “But we don’t understand how most of it works.”

Hughes and his colleagues used a high-throughput method to identify short strings of RNA letters, called ‘motifs,’ that RBPs recognize. On a long sequence of RNA, motifs act as flags that show RBPs where to bind.

Hughes’s method uncovered these binding sites under circumscribed laboratory conditions, however, so it’s unclear whether the data will hold true in living systems, experts note.

Still, some autism researchers are keen to use the database, dubbed the Catalog of Inferred

Sequence Binding Preferences of RNA, or **CISBP-RNA**. It could be used to find binding sites within **genetic sequences of individuals with autism**, for example, or to learn where a known RBP binds.

"It's a remarkable resource," says **Genevieve Konopka**, assistant professor of neuroscience at the University of Texas Southwestern Medical Center in Dallas, who was not involved in the study.

Konopka's group and others are investigating how several RBPs are involved in neuronal function and brain disorders. "It's going to be great to mine all this data that they've provided," she says.

## RNA pools:

Most researchers trying to find the binding sites of RBPs use a 'cross-linking' approach. They first freeze RNA-protein interactions within living cells, purify the protein of interest, and then sequence the bound RNA letter by letter.

That approach "has a lot of physiologic relevance, but it's really laborious," notes **Jennifer Darnell**, research associate professor of molecular neuro-oncology at Rockefeller University in New York City, whose lab has used this method to **identify FMRP's binding sites**.

In contrast, Hughes' method places a single RBP into a pool of some 240,000 pieces of synthetic RNA, purposely selected to represent the widest possible variety of sequences. "Then we just measure what stuck to the protein and what didn't," Hughes says.

When Hughes debuted the method, called RNAcompete, in 2009, he reported the binding preferences of nine RBPs<sup>3</sup>. In the new study, he identified motifs for 207 proteins: 85 from humans, 61 from fruit flies and 61 from 18 other eukaryotic species.

Many proteins are evolutionarily conserved — that is, they have remained the same or close to it over the course of evolution across species. The database assumes that if two proteins — from any species — are at least 70 percent identical, then they will have the same RNA-binding motif. Based on that reasoning, the database holds motifs of more than half of the thousands of known human RBPs, the study found.

This resource complements another collection of RNA-binding proteins in a database called **RBPDB** (the database of RNA-binding protein specificities), which Hughes' team created in 2010 for about 70 other proteins reported in published studies. Hughes says he plans to combine the two resources eventually.

Autism researchers say they are eager to try the new database, though they point out that the data for any particular RBP will need to be verified *in vivo*.

"This is a very systematic resource, but I think it's only the first step," says **Peng Jin**, professor of human genetics at Emory University in Atlanta, who was not involved in the new work. Jin points out that the RNA interactions of FMRP and other RBPs often depend on the particular cellular environment.

Perhaps because of this, the binding sites the new method identified for FMRP do not match those reported in previous studies. "I'm not sure what to exactly make of this data on FMRP," Darnell says. "Basically, what they pulled out doesn't match what anybody else has found."

Still, the methods seem to work well for some RBPs. Darnell and others say they are particularly intrigued by the findings for RBFOX1. Previous work has found that the RBFOX1 protein is expressed at **abnormally low levels** in postmortem autism brains.

Last year, Konopka and her colleagues showed that knocking down the levels of RBFOX1 in cell culture **affects the splicing** of hundreds of genes<sup>4</sup>. Splicing is a process in which strings of RNA are shuffled around, allowing a single gene to encode multiple proteins.

In the new study, Hughes found that a shortage of RBFOX1 can also cause RNA to degrade before it is turned into protein. Using published datasets of gene expression in postmortem brain tissue, his team found that individuals with low levels of RBFOX1 also tend to have too little of the RNAs that encode proteins in potassium ion channels. These molecular conduits control the flow of electrical current between connected neurons.

"So there you can start to build a story that perhaps RBFOX1 is playing a role in neuronal ion channel function and connectivity," Konopka says. Intriguingly, FMRP has also been **linked to these channels**.

"The ultimate goal is moving toward understanding what part of cellular function these genes are involved in," Konopka adds. "And then we can try to identify drugs that affect those functions."

## References:

1. Ray D. *et al. Nature* **499**, 172-177 (2013) [PubMed](#)
2. Sebat J. *et al. Science* **316**, 445-449 (2007) [PubMed](#)
3. Ray D. *et al. Nat. Biotechnol.* **27**, 667-670 (2009) [PubMed](#)
4. Fogel B.L. *et al. Hum. Mol. Genet.* **21**, 4171-4186 (2012) [PubMed](#)