Scientists trace extensive networks regulating alternative RNA splicing

September 20 2008

RNA targets of tissue-specific splicing factors Fox-1 and Fox-2 are successfully predicted

Two professors at Cold Spring Harbor Laboratory (CSHL) have succeeded in tracing intricate biochemical networks involving a class of proteins that enable genes to express themselves in specific tissues at particular moments in development.

Michael Q. Zhang, Ph.D. and Adrian R. Krainer, Ph.D., both professors at CSHL and heads of laboratories, are exploring a phenomenon that biologists and geneticists call RNA splicing. Splicing is a key step in the multi-step process that transmits a gene's instructions to a cell, telling it how and when to manufacture specific protein molecules, and how much to produce.

Poorly understood until recently, the splicing machinery and the networks that control it are only now coming into clear view. In a paper appearing this week in the journal *Genes & Development*, Drs. Zhang, Krainer, and colleagues from CSHL, Stony Brook University, and Rosetta Inpharmatics, reveal how two closely related proteins called Fox-1 and Fox-2--which are two among many splicing factors--control regulatory networks involving many other genes.

These regulatory networks, which are surprisingly extensive and highly conserved by evolution, help scientists gain insights into gene regulation in different cells—in these experiments, brain and muscle cells. The work is also relevant to understanding dysfunction, which in brain and

muscle has been implicated in a range of developmental illnesses from autism to heart disease.

Alternative splicing helps explain human complexity

Fox-1 and Fox-2 are two among several hundred splicing factors, a class of proteins whose highly specialized functions help explain human complexity. Biologists involved in the Human Genome Project were frankly astonished to discover that everything that makes us human is the product of a set of only 23,000 or so genes. That number in itself, though several times smaller than prior estimates, is not shocking; it is the relative size of other genomes that surprised scientists.

The common fruit fly that hovers over your ripening bananas, for instance, possesses some 14,000 genes. It's perfectly obvious that human beings are vastly more complex, biologically, than a fly. Molecular biologists have demonstrated in recent years that it is not the number of genes that is the key to complexity but rather the number and diversity of gene products that a given set of genes can instruct cells to manufacture.

Rather than a single gene ordering the production of a single kind of protein, as scientists used to assume, it turns out that individual genes can in some cases give rise to dozens or even thousands of different proteins, thanks to a phenomenon called alternative splicing.

Zhang and Krainer's new research focused on splicing, which is the "editing" of RNA molecules generated by activated genes. By cutting up and pasting back together bits of these RNA intermediaries, the splicing machinery deletes "non-coding" segments (called introns) and stitches together "coding" segments (called exons).

The final product of RNA splicing (called a mature messenger-RNA

"transcript") finds its way out of the cell nucleus. Carrying the "message" of an activated gene, it enters a structure called the ribosome, which reads its coding instructions and manufactures a protein with a particular configuration. The protein's shape--and function--will vary depending on how the splicing factors back in the nucleus have cut and pasted together the final RNA transcript which has served as the blueprint for the protein's manufacture.

How are splicing-factor targets recognized?

The question addressed by Zhang and Krainer is how particular splicing factors can recognize their specific targets. How do these proteins know where to attach to raw, "unedited" RNA transcripts, and how do they engage the cellular machinery that actually splices RNAs--a complex called the splicesosome? What accounts for differing "edits" of the same RNA molecule under the direction of a given splicing factor? What can go wrong in splicing and how can such anomalies result in cellular dysfunctions that culminate in human diseases?

In their newly published research, Zhang and Krainer used a comprehensive set of computational methods to predict how and where the splicing factors Fox-1 and Fox-2 attach, or bind, to unedited target RNA transcripts. "The effect of splicing factors on the activation or repression of splicing often depends on the location and context of the RNA sequences they bind," Dr. Krainer explained.

It was known that the two proteins have the ability to home in on a stretch of RNA that bears a specific code ("UGCAUG"), although the presence of this sequence on an RNA is not sufficient to guarantee binding by Fox-1/2 proteins. "Prior to our experiments, only a handful of targets of this kind had been determined experimentally," Dr. Zhang observed. An authority on the application of computational methods to biological questions, Zhang and members of his lab generated

predictions about Fox-1 and Fox-2 after closely examining the genomes of 28 different vertebrate species.

From conserved networks to disease targets

Experimental follow-up in the Krainer lab of the predicted binding sites indicated that between one-half and three-fourths of the thousands of computational binding-site predictions—and the corresponding alternative splicing patterns—were correct. The fact that the predicted targets were conserved during evolution across distantly related species suggests the relative importance of Fox-1 and Fox-2 function in many living systems. Indeed, the experiments confirmed that the regulatory networks that govern the actions of the splicing factors are also conserved across species.

Perhaps most important, the team found, in Zhang's words, that "many of the predicted RNA targets play important roles in neuromuscular functions and disorders." That is a direct reflection of the fact, also demonstrated in the experiments, that Fox-1 and Fox-2 are either exclusively or preferentially expressed in brain, heart, and skeletal muscle. This is consistent with prior studies showing that some people with genetic diseases including epilepsy, mental retardation and autism, have mutations of the Fox-1 gene or abnormally express it.

"The Fox-1 and -2 proteins are essential regulators of specific splicing events that occur in cells of particular tissues," Krainer said. "Our ability to trace their regulatory networks and observe how they are conserved, albeit with significant variation, across and within species, may lead to potential approaches for intervention in cases where, because of genetic mutation, vital cellular functions are perturbed. One key to this is our ability to place discrete disease-associated genes into well-defined gene-regulatory networks. We anticipate that our continuing work will shed new light on such networks."

Citation: "Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2" appears in the September 2008 issue of Genes & Development. The full citation is: Chaolin Zhang, Zuo Zhang, John Castle, Shuying Sun, Jason Johnson, Adrian R. Krainer and Michael Q. Zhang. The paper can be obtained online: doi:10.1101/gad.1703108.

Source: Cold Spring Harbor Laboratory