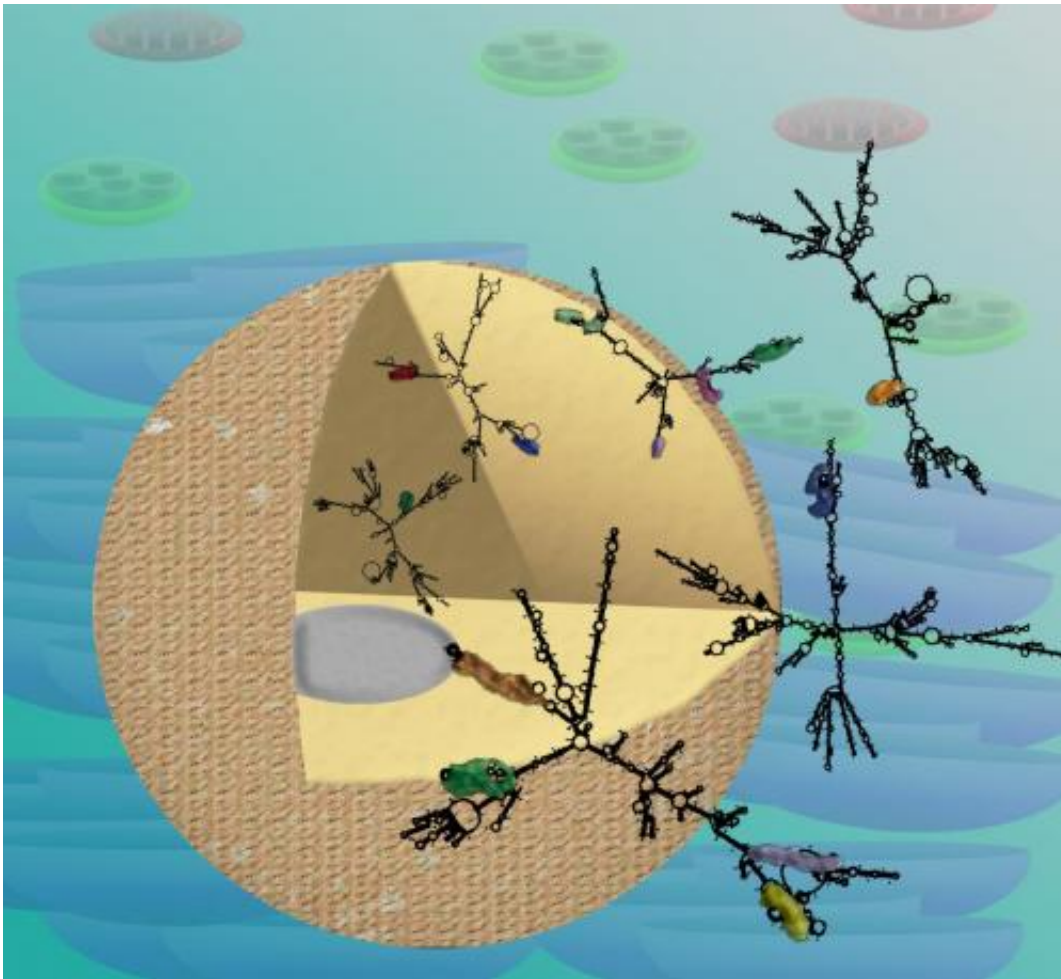


Scientists identify patterns of RNA regulation in the nuclei of plants

December 31 2014



In a new study done in plants, University of Pennsylvania biologists give a global view of the patterns that can affect the various RNA regulatory processes that occur before these molecules move into the cytoplasm, where they are translated into the proteins that make up a living organism. Credit: University of Pennsylvania

When the human genome was first sequenced, experts predicted they would find about 100,000 genes. The actual number has turned out to be closer to 20,000, just a few thousand more than fruit flies have. The question logically arose: how can a relatively small number of genes lay the blueprint for the complexities of the human body?

The explanation is that genes are subject to many and varied forms of regulation that can alter the form of that protein and can determine whether and how much of a [gene product](#) is made. Much of this regulation occurs during and just after DNA is transcribed into RNA.

In a new study done in plants, University of Pennsylvania biologists built on earlier work in which they cataloged all the interactions that occur between RNA and the proteins that bind to it. This time, they looked exclusively at these interactions in the nuclei, and simultaneously obtained data about the nuclear RNA molecules' structure. By combining these datasets, their findings give a global view of the patterns that can affect the various RNA regulatory processes that occur before these molecules move into the cytoplasm, where they are translated into the proteins that make up a living organism.

In addition, the researchers have provided a vast, publically available set of data that other scientists can use to address questions about any genes and regulatory mechanisms that interest them, gaining a better understanding of the dynamics of the journey from DNA to protein.

Brian D. Gregory, an assistant professor in Penn's School of Arts & Sciences' Department of Biology, was senior author on the work, which will appear in the journal *Molecular Cell*. Sager J. Gosai, a research specialist, and Shawn W. Foley, a graduate student, both members of Gregory's lab, were co-first authors. Additional contributors from Penn included Ian M. Silverman, a graduate student in the Gregory lab, along with Fevzi Daldal, a professor in the Department of Biology and Nur

Selamoglu of the Daldal lab. The Penn researchers teamed with Emory University's Dongxue Wang and Roger B. Deal and University of Arizona's Andrew D. L. Nelson and Mark A. Beilstein to conduct the study.

Earlier this year in *Genome Biology*, Gregory's team reported on a method they developed to obtain a complete catalog of the interactions in live organisms between RNA and RNA-binding proteins, or RBPs, which interact with RNA transcripts to repress, enhance or otherwise alter gene expression in a cell-type specific manner. The technique is called PIP-seq, for protein interaction profile sequencing. Their initial demonstration of PIP-seq identified the full complement of RBP interaction sites in a human cell line.

In the current work, they used the commonly studied plant *Arabidopsis thaliana* to map out all of the RBP interaction sites as well as compile a full look at the secondary structure of the RNA transcripts. Unlike the first study, which looked at all the RNA in the cell, a set of material known as the transcriptome, this study looked only in the nucleus.

"By focusing specifically on the nucleus we can get away from all of the features on RNA molecules that are associated with the process of translation into proteins, which occurs in the cytoplasm," Gregory said.

The researchers extracted nuclei from 10-day-old *Arabidopsis* seedlings. They performed PIP-seq and also obtained information on the secondary structure of the RNA—how the strands of RNA fold, loop or bind together.

Focusing on sections of RNA that bind to RBPs, the team found that these sequences have been conserved over evolutionary time and are likely playing an important function in gene regulatory mechanisms.

The scientists also found a strong inverse relationship between patterns of RBP binding and secondary structure.

"When structure is low, proteins tend to bind those regions and when structure is high, RBPs tend to not bind those regions," Gregory said. "Time and time again, we've seen that the structural context, and not just the RNA sequence, is a selective force in RBP binding."

Another significant finding was unique patterns of RBP binding and structure present around the start codon of each messenger RNA transcript, which is where a cell's protein-making machinery begins the process of making RNA in proteins.

"This is suggesting that there is a regulatory event happening here even before the RNA comes out of the nucleus and engages with the translation machinery," Gosai said. "It's an exciting place for future studies to start with and figure out what regulation events are happening in the nucleus."

Two key forms of transcript regulation are [alternative splicing](#), in which pieces of RNA undergo a cut-and-paste process to generate new sequences that can code for various proteins, and alternative polyadenylation, which alters where a transcript ends and an adenine "tail" is added, a process that can enhance either stabilization or degradation of the RNA molecule.

In their analysis, the Penn biologists found that RBP-binding sites and certain patterns of secondary structure were much more common at sites where alternative splicing and alternative polyadenylation occur.

"In humans, almost 95 percent of genes are alternatively spliced, and the number is at least 60 percent in plants," said Foley. "To see high levels of RBP binding and an interplay with secondary structure at sites of

alternative splicing and polyadenylation in plants is good indication of where and how regulation is occurring to produce different proteins from one RNA sequence."

As in their previous study using PIP-seq, Gregory and his colleagues identified recurring patterns, known as "motifs," of RNA sequences at sites that tended to be bound by certain RBPs. It's possible, the researchers noted, that these groups of RBPs could bind functionally-related [genes](#) to coordinate their regulation.

Finally, the team zoomed in on one RBP-bound sequence motif that was particularly abundant and found that it interacted with an RBP called CP29A.

"This protein was known to bind RNA in the chloroplast, but we were able to identify it as a nuclear RBP for the first time," Foley said, suggesting CP29A may be an important regulatory factor in both organelles.

To follow up on this work, the Penn scientists will examine how RNA regulation differs in plant tissues at different developmental stages. They also plan to use PIP-seq and structural analyses to study other types of organisms.

"Now that we've found beautiful patterns that mark alternative splicing and other events that shape the protein-coding capacity of plants, we're going to go in and identify the proteins that lead to those," Gregory said. "And eventually we'd like to go into humans and other organisms and ask if we see similar patterns."

Provided by University of Pennsylvania

Citation: Scientists identify patterns of RNA regulation in the nuclei of plants (2014, December 31) retrieved 19 April 2024 from <https://phys.org/news/2014-12-scientists-patterns-rna-nuclei.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.