

# Research reveals how cells process large genes

August 22 2005

---

Important messages require accurate transmission. Big genes are especially challenging because they combine many coding segments (exons) that lie between long stretches of non-coding elements (introns). During processing, introns are snipped out and exons pasted together to form a template for proteins called messenger RNA (mRNA). Mistakes in RNA processing can reduce the expression of a functional protein or, worse, produce an abnormal protein that interferes with normal cell behavior. But just how a cell's molecular machinery eliminates long introns without making errors has puzzled scientists for years.

Now, investigators at Carnegie Mellon University have discovered that a novel mechanism, called recursive splicing, removes long introns by steadily paring them down in a predictable fashion and joining the remaining exons. The findings are published this summer in *Genetics*. This process, which the investigators discovered in the fruit fly *Drosophila*, has been conserved over tens of millions of years of insect evolution and also appears likely to occur in humans, according to the investigators.

"While some scientists have suspected that large introns might not be removed in one piece through direct splicing, no one had identified how this could happen. Now we have identified a way," said Antonio-Javier Lopez, professor of biological sciences at Carnegie Mellon. Ultimately, recursive splicing could be responsible for thwarting molecular mishaps in the expression of large human genes associated with diseases like muscular dystrophy, cystic fibrosis and cancer.

"We found that many large introns are removed by multiple recursive splicing steps," Lopez said. "These steps involve the sequential excision of smaller subfragments. Our work also indicates that most recursive splicing events leave no clues in the final mRNA. This is why they have not been detected before now."

These previously undetected events could have profound implications for predicting what constitutes a gene and for studying gene expression, mutation and evolution, according to Lopez.

For example, recursive splicing must now be taken into account when evaluating mutations that disrupt gene expression and produce a dysfunctional or non-functional protein.

"Current data indicate that at least 15 percent of disease-causing mutations occur at standard signals where intron removal takes place through direct splicing. Mutations at recursive splice sites may cause additional diseases, but until now we haven't looked for them."

Knowledge of recursive splicing also will help investigators predict structures of genes that span large intervals of DNA, Lopez said.

Recursive splicing relies on the unusual activity of a ratchetting point, a pattern of chemical groups (nucleotides) previously discovered within the genome by Lopez. One end of a ratchetting point contains a sequence of nucleotides similar to the signal normally found at the beginning of an intron. This signal is juxtaposed with another sequence like that normally found at the end of an intron. Such a unique pairing allows a ratchetting point to function sequentially as an acceptor for splicing to an upstream exon and then as a donor for splicing to the next downstream ratchetting point or exon. As the process goes from ratchetting point to ratchetting point, small signature loops of RNA called lariats are released from the intron. Repeated over and over, recursive splicing eventually binds, or

ligates, two distant exons.

Lopez's team developed molecular tools to analyze the lariats released from any intron during splicing in vivo. In his analyses, he found that the production of recursive lariats greatly exceeded that of direct lariats, indicating that recursive splicing is the predominant processing pathway for long introns. Lopez combined these experimental data with computational and phylogenetic analyses of several fruitfly and other insect species.

"Our experimental results agreed with the computational findings, indicating that these ratchetting points mediate the removal of intron subfragments in one direction as the gene is transcribed initially from DNA into RNA," Lopez said.

Lopez found that predicted recursive splice sites were 10 times more likely than expected to be found in introns greater than 200 kilobases in length, and 92 percent of them were conserved over at least 25 million years of insect evolution. This discovery strongly suggests that recursive splicing plays a special role in the correct expression of large genes. Bioinformatic and phylogenetic analyses conducted by Lopez also indicate that recursive splicing plays a role in at least 124 fruitfly introns, with up to seven potential cutting steps identified for a single intron. Similar analyses suggest that the same process also occurs in large introns of mammals, including humans, and the Lopez team is now testing this hypothesis experimentally.

"The striking evolutionary conservation of ratchetting points suggests that recursive splicing provides specific advantages for large introns," Lopez said. "One possibility is that recursive splicing prevents the generation of long RNA transcripts that could form structures that interfere with correct processing into mRNA. Another is that recursive splicing might help stimulate transcription through long introns by

promoting interactions between the splicing and transcription machineries. We also know already that recursive splicing is used to control the removal of certain exons from mRNAs, generating structural and functional variation among the gene products."

Source: Carnegie Mellon University

Citation: Research reveals how cells process large genes (2005, August 22) retrieved 24 May 2024 from <https://phys.org/news/2005-08-reveals-cells-large-genes.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.