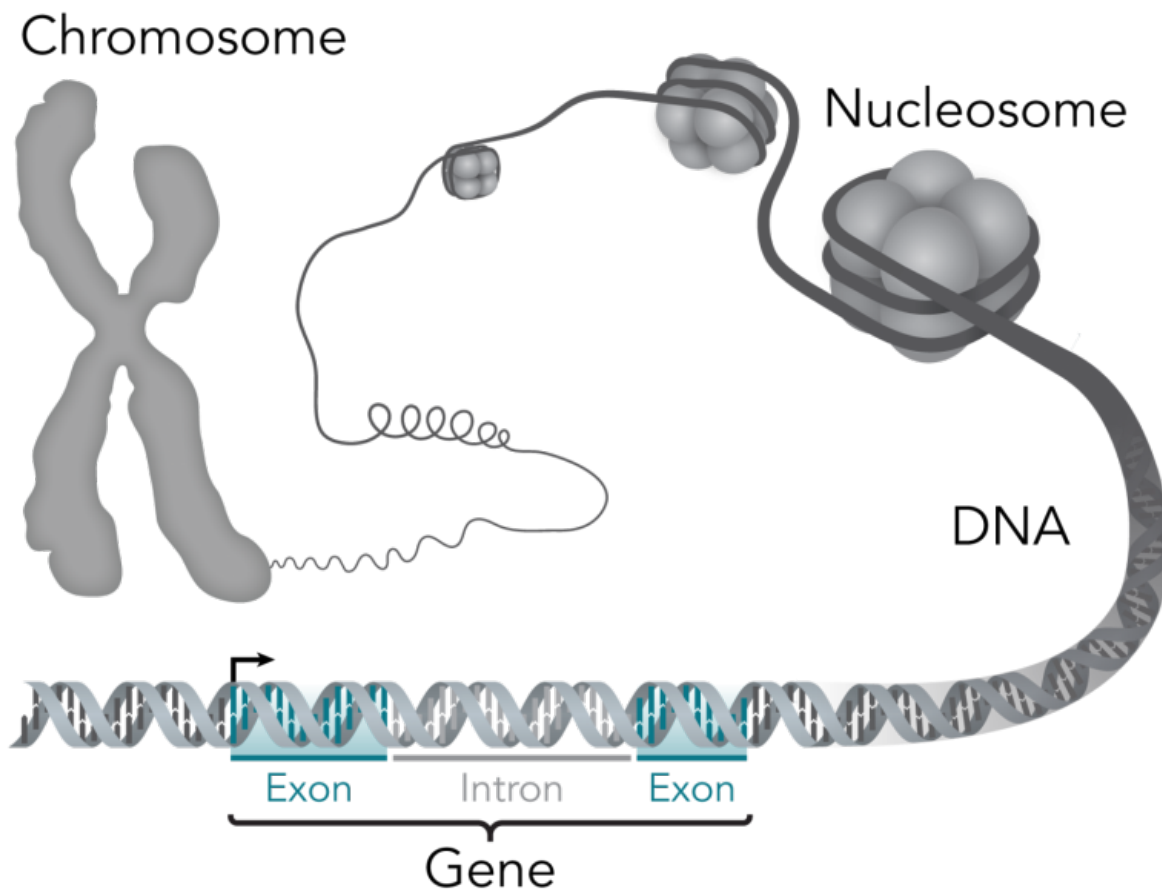


Alternative proteins encoded by the same gene have widely divergent functions in cells

February 11 2016



This stylistic diagram shows a gene in relation to the double helix structure of DNA and to a chromosome (right). The chromosome is X-shaped because it is dividing. Introns are regions often found in eukaryote genes that are removed in the splicing process (after the DNA is transcribed into RNA): Only the exons encode the protein. The diagram labels a region of only 55 or so bases as a gene. In reality, most genes are hundreds of times longer. Credit: Thomas

It's not unusual for siblings to seem more dissimilar than similar: one becoming a florist, for example, another becoming a flutist, and another becoming a physicist.

Something of the same diversity applies to the "brood" of proteins produced from any single gene in human cells, a new study led by scientists at Dana-Farber Cancer Institute, University of California, San Diego School of Medicine, and McGill University has found. In a first large-scale systematic study, the researchers found that most sibling proteins—known as "protein isoforms" encoded by the same gene—often play radically different roles within tissues and cells, however alike they may be structurally.

The research, published online today by the journal *Cell*, stands to have a powerful effect on the understanding of human biology and the direction of future research. For one, it may help explain how the mere 20,000 protein-coding [genes](#) in the human genome - fewer than are found in the genome of a grape—can give rise to creatures of such enormous complexity. Scientists know that the number of different proteins in human cells, thought to be upwards of 100,000, far exceeds the number of genes, but many questions have remained. Do most of those proteins have a unique function in the cell, or do their roles sometimes overlap? The discovery that different protein isoforms encoded by the same gene may have divergent functions on a larger scale than realized suggests that they vastly multiply what our genes are capable of.

This diversity also suggests that each protein isoform needs to be studied individually to understand its normal role and its potential involvement in disease, the study authors state.

"Research into cancer-related proteins, for example, often focuses on the most prevalent isoforms in a given cell, tissue, or organ," said co-senior author David E. Hill, PhD, associate director of the Center for Cancer Systems Biology (CCSB) at Dana-Farber. "Since less-prevalent protein isoforms may also contribute to disease, and may prove to be valuable targets for drug therapy, their role should be examined as well; and to do that properly, we also need comprehensive clone collections covering all expressed isoforms."

Previous functional studies of protein isoforms have generally been done on a gene-by-gene basis. Furthermore, researchers frequently compared the activity of a gene's "minor" isoforms to that of its predominant isoform in a particular tissue. The new study approached the functional question from a larger perspective - by gathering multiple protein isoforms of hundreds of genes and comparing how they specifically interact with any other [human protein](#).

One of the ways that cells produce multiple protein isoforms from individual genes is a process called alternative splicing. Most human genes contain multiple segments called exons, separated by intervening non-coding sequences called introns. In the cell, different combinations of these individual exons are "glued" or spliced together to generate a final expressed gene product; thus, a single gene can encode a set of distinct, but related protein isoforms, depending on the specific exons that are spliced. One isoform, for example, may result from splicing exons A-B-C-D of a particular gene. Another may arise from the skipping of exon C, resulting in a product with only exons A-B-D.

For the new study, researchers devised a technique called "ORF-Seq" that allowed them to identify and clone large numbers of alternatively spliced gene products in the form of open reading frames (ORFs), and use them to produce multiple protein isoforms for hundreds of genes.

Of the roughly 20,000 genes in the human genome that code for proteins, researchers concentrated on about eight percent. Using ORF-Seq, they ultimately created a collection of 1,423 protein isoforms for 506 genes, of which more than 50 percent were entirely novel gene products. They subjected 1,035 of these protein isoforms through a mass screening test that paired them with 15,000 human proteins to see which would interact.

"The exciting discovery was that isoforms coming from the same gene often interacted with different protein partners," remarked Gloria Sheynkman, PhD, of Dana-Farber and one of the lead authors. "This suggests that the isoforms play very different roles within the cell" - much as siblings with different careers often interact with different sets of friends and co-workers.

The researchers found that in most cases, related isoforms shared less than half of their protein partners. Sixteen percent of related isoforms share absolutely no protein partners. "From the perspective of all the protein interactions within a cell, related isoforms behave more like distinct proteins than minor variants of one another," Tong Hao, of Dana-Farber and one of the lead authors, asserted.

Intriguingly, isoforms that stem from a minuscule difference in DNA - a difference of just one letter of the genetic code—sometimes had starkly different roles within the cell, researchers found. At the same time, related isoforms that are structurally quite different may have very similar roles.

Quite often, the interaction partners of related isoforms vary from tissue to tissue, the researchers found. In the liver, for example, an isoform may interact with one set of proteins. In the brain, a relative of that isoform may interact with a largely different set of protein partners.

"A more detailed view at [protein](#) interaction networks, as presented in our paper, is especially important in relation to human diseases," said co-senior author Lilia Iakoucheva of UC San Diego. "Drastic differences in interaction partners among splicing isoforms strongly suggest that identification of the disease-relevant pathways at the gene level is not sufficient. This is because different variants could participate in different pathways leading to the same disease or even to different diseases. It's time to take a deeper dive into the networks that we are building and analyzing."

Provided by Dana-Farber Cancer Institute

Citation: Alternative proteins encoded by the same gene have widely divergent functions in cells (2016, February 11) retrieved 26 April 2024 from <https://medicalxpress.com/news/2016-02-alternative-proteins-encoded-gene-widely.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.