

February 11, 2016 | By Heather Buschman, PhD and Anne Doerr, Dana-Farber Cancer Institute

Alternative Proteins Encoded by Same Gene Have Widely Divergent Functions

Research helps explain complexity of species that have fewer than expected genes, and highlights the need to study all related protein forms for their potential roles in disease

A single human gene can produce many different proteins. In the first large-scale study of its kind, researchers at University of California, San Diego School of Medicine, Dana-Farber Cancer Institute and McGill University report that most of these sibling proteins encoded by the same gene — known as protein isoforms — often play radically different roles within tissues and cells.

The study, published February 11 by *Cell*, stands to have a powerful effect on the understanding of human biology and may influence future research in many fields, the researchers said.

For example, the study may help explain how 20,000 protein-coding genes in the human genome — fewer than found in the genome of a grape — can give rise to creatures of such enormous complexity. This diversity in protein function suggests that each protein isoform should be studied individually to understand its normal role and its potential involvement in disease, the researchers said.

“Research into cancer-related proteins, for example, often focuses on the most prevalent isoforms in a given cell, tissue or organ,” said David Hill, PhD, of 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 collections covering all expressed isoforms.” Hill was a co-senior author of the study with Lilia Iakoucheva, PhD, assistant professor of psychiatry at UC San Diego School of Medicine.

Previous studies of protein isoforms have generally been conducted on a gene-by-gene basis. Researchers frequently compared the activity of a gene’s “minor” isoforms to that of its predominant isoform in a particular tissue.

The new study sought a larger perspective by gathering multiple protein isoforms encoded by hundreds of genes and comparing how they specifically interact with any other human protein. Of the roughly 20,000 genes in the human genome that code for proteins, the researchers concentrated on about eight percent. Using a new technique they devised called “ORF-Seq,” the team 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 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 protein interactions within a cell, related isoforms behave more like distinct proteins than minor variants of one another, the researchers found.

The team also discovered that protein isoforms stemming from a minuscule difference in DNA — a difference of just one letter of the genetic code — sometimes had starkly different roles within the cell. At the same time, other related isoforms that are structurally quite different may have very similar roles.

Quite often, the interaction partners of related isoforms varied 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.

“This detailed view of protein interaction networks is especially important in relation to human diseases,” Iakoucheva said. “Drastic differences in interaction partners among protein 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.”

Co-authors of the study also include Shuli Kang, UC San Diego; Xinping Yang, Dana-Farber, Harvard Medical School and Nanfang Hospital, Southern Medical University; Jasmin Coulombe-Huntington, McGill University; Gloria Sheynkman, Tong Hao, Aaron Richardson, Yun Shen, Ryan Murray, Kerstin Spirohn, Bridget Begg, Andrew MacWilliams, Quan Zhong, Shelly Trigg, Stanley Tam, Lila Ghamsari, Nidhi Sahni, Song Yi, Maria Rodriguez, Dawit Balcha, Kourosch Salehi-Ashtiani, Benoit Charloteaux, Alyce Chen, Michael Calderwood, Marc Vidal, Dana-Farber and Harvard Medical School; Song Sun, University of Toronto, Mt. Sinai Hospital and Uppsala University; Fan Yang, University of Toronto and Mt. Sinai Hospital; Miquel Duran-Frigola,

Barcelona Institute of Science and Technology; Samuel Pevzner, Dana-Farber and Boston University; Guihong Tan, Michael Costanzo, Brenda Andrews, Charles Boone, University of Toronto; Xianghong Zhou, University of Southern California; Patrick Aloy, Barcelona Institute of Science and Technology and Institució Catalana de Recerca i Estudis Avançats; Frederick Roth, Dana-Farber, University of Toronto, Mt. Sinai Hospital and Canadian Institute for Advanced Research; Yu Xia, Dana-Farber and McGill University.

MEDIA CONTACT

Heather Buschman, 858-249-0456, hbuschman@ucsd.edu

UC San Diego's [Studio Ten 300](#) offers radio and television connections for media interviews with our faculty, which can be coordinated via studio@ucsd.edu. To connect with a UC San Diego faculty expert on relevant issues and trending news stories, visit <https://ucsdnews.ucsd.edu/media-resources/faculty-experts>.