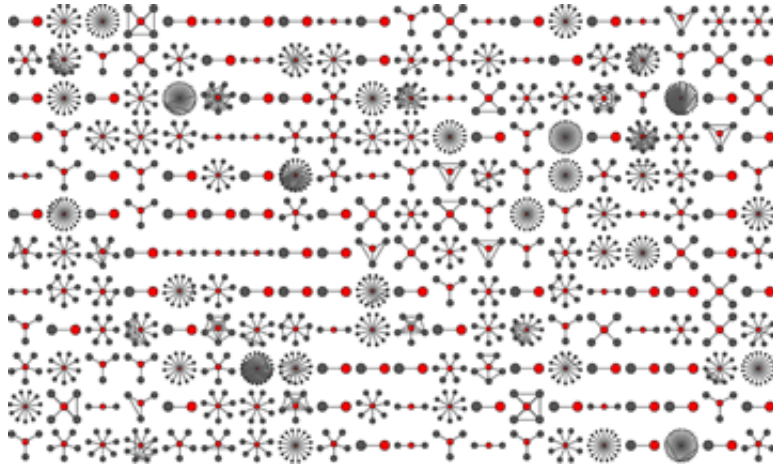


# Facebook for the proteome

July 17 2015, by Elizabeth Cooney

---



Mapping the interactome helps characterize unknown proteins. Credit: Harper and Gygi labs

There are approximately 20,000 human genes that encode proteins, but despite remarkable progress since the human genome was first sequenced more than a decade ago, scientists still understand in detail how only a small fraction of how these proteins function in the cell.

The field of proteomics, a logical extension of the Human Genome Project, is revealing how proteins execute processes encoded by the genes that produce them.

Advanced tools have identified proteins—and the networks of partners they work with—that drive cellular function.

Focusing on particular families of proteins, cell biologists have parsed how they interact in common pathways to influence health and disease.

Now Harvard Medical School scientists are telling the much larger story of the complexes and clusters of related proteins that make up what is known as the human "interactome."

A team led by Wade Harper and Steven Gygi has begun to capture and catalog all [protein](#) complexes in the [human proteome](#), creating a map they call the BioPlex network. Harper is the HMS Bert and Natalie Vallee Professor of Molecular Pathology and chair of the Department of Cell Biology; Gygi is an HMS professor of cell biology and director of the Thermo Fisher Center for Multiplexed Proteomics at Harvard Medical School.

## First fruits

Reporting in the journal *Cell*, they describe the first fruits of their systematic exploration of the interactome. Using high-throughput affinity purification paired with mass spectrometry, they applied their approach to more than 10 percent of the human proteome. In affinity purification, a single protein is used as "bait" to capture interacting partners. Mass spectrometry identifies proteins by measuring their mass and other attributes.

"You could say we are creating a Facebook for the human proteome," said Edward Huttlin, research fellow in the Gygi lab and lead author of the paper. "We are trying to figure out which proteins are related to which other proteins in the cell in the same way we might learn about a person by asking who their friends are and who they associate with."

In the largest study of its kind, they identified nearly 24,000 interactions among more than 7,600 proteins—86 percent of them previously

unknown. The network's architecture predicts where proteins are located within the cell, what biological processes they are part of and what molecular roles they may play.

By looking at these interaction patterns and the "communities" that exist within the network, Huttlin said, scientists are able to develop hypotheses concerning the potential functions of poorly studied components of the network.

There are other databases that describe protein interactions in mammalian cells, but these are incomplete. The goal of the HMS team is to comprehensively map individual protein functions as well as chart the organization of the entire proteome.

## **Challenges of scale**

Working at such a large scale introduces numerous challenges, including maintaining pipeline quality, keeping the instrumentation running, and developing ways to maintain and organize large quantities of proteomic data. This infrastructure is funded by Biogen and the National Human Genome Research Institute, and the data obtained are deposited in a public resource called [BioGRID](#) and in the [BioPlex website](#) hosted at HMS.

The data are made freely available through these databases prior to publication—a feature that distinguishes this approach from other efforts of its kind.

Beyond scale, the team's approach offers another advantage, Huttlin said. They look at protein interactions within an intact cell as opposed to seeing the proteins in isolation, which may explain in part why they have found new connections that the act of separation might destroy.

## ALS proteins

To demonstrate the power of their work, they examined genetic mutations and related proteins involved in Amyotrophic Lateral Sclerosis, also known as Lou Gehrig's disease. They found previously unsuspected relationships among proteins expressed by genetic sequence variations found in people with the disease.

Work continues on the other 90 percent of the human proteome, Harper said. Indeed, as of July 1, more than 50,000 interactions derived from experiments on more than 25 percent of [human](#) genes have been deposited into BioGRID through the efforts of the HMS team. He expects this resource to generate hypotheses for countless future studies in cell biology now that the BioPlex network has trained its light on what is considered the "dark matter" of the proteome.

"We wanted to be able to describe as many molecular machines in the cell as possible—those complexes that interact," said Harper. "We've partially succeeded at that, and it's already a large amount of data compared to what's been done before. But there is a lot more work to be done."

## Secrets

"For me, the most astounding part of this enterprise has been how easily many [human genes](#) with completely unknown function give up their secrets," Gygi said. "We just had to express the protein, and through its associations one can often accurately infer that gene's purpose in life."

"We consider the ongoing Biogen collaboration with the Harper and Gygi laboratories at HMS on protein interactions to be very significant and central in our scientific efforts to explore the underlying

mechanisms of disease," said Spyros Artavanis-Tsakonas, chief scientific officer at Biogen, HMS professor of [cell biology](#), emeritus, and a co-author on the study. "By better understanding the way proteins function and interact, we hope to gain new insights into complex neurodegenerative diseases like Alzheimer's, ALS and Parkinson's that may one day lead to important new treatments."

## Work in progress

Like Facebook, the interactome will keep growing.

"There are so many leads here at this point, there's no way we can pursue them all," Huttlin said. "It'll be really fun to see what other people can find as they go through and build upon this network."

**More information:** "The BioPlex Network: A Systematic Exploration of the Human Interactome." DOI: [dx.doi.org/10.1016/j.cell.2015.06.043](https://doi.org/10.1016/j.cell.2015.06.043)

Provided by Harvard Medical School

Citation: Facebook for the proteome (2015, July 17) retrieved 21 June 2024 from <https://phys.org/news/2015-07-facebook-proteome.html>

<p>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.</p>
--