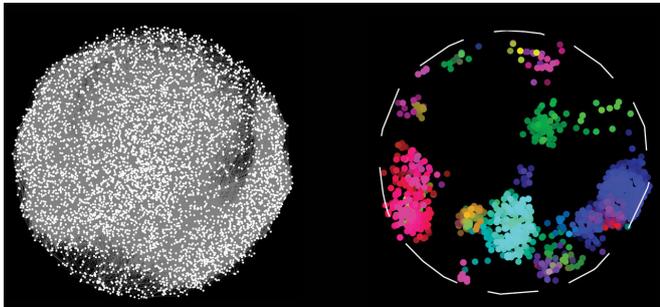# World's largest map of protein connections holds clues to health and disease

8 April 2020, by Jovana Drinjakovic

By charting pairwise interactions between 17,500 human proteins, scientists have created a map, on the left, depicting which proteins work together to sustain cellular function. Proteins with similar interactions profiles fall into discrete colour-coded clusters representing different bioprocesses in the cell. Credit: Katja Luck et al.

The human body is composed of billions of cells, each of which is made and maintained through countless interactions among its molecular parts. But which interactions sustain health and which ones can cause disease when they go awry? The human genome project has provided us with a "parts list" for the cell, but only if we can understand how these parts go together, or interact, can we really begin to understand how the cell works and what goes wrong in disease.

To answer these questions, scientists needed a reference map of interactions—an interactome—between gene-encoded proteins, which make up cells and do most of the work in them.

"Since the mid-1990s, our collaborative team has pushed the idea that interactome maps can illuminate fundamental aspects of life," says Marc Vidal, one team leader and Director of the Center for Cancer Systems Biology (CCSB) at Dana-Farber Cancer Institute in Boston.

"Our paper describes the first human interactome

reference map, constituting a "scaffold" of information to better understand how faulty genes cause diseases such as cancer, but also how viruses such as the coronavirus that causes COVID-19 interact with their host human proteins," says Vidal.

Almost a decade in the making, the human protein map is now available thanks to a joint effort, involving over 80 researchers in the United States, Canada, Spain, Belgium, France and Israel, jointly led by Vidal, David E Hill and Michael A Calderwood, at Dana-Farber Cancer Institute, Frederick P Roth, at the University of Toronto's Donnelly Centre for Cellular and Biomolecular Research.

The largest of its kind, the Human Reference Interactome (HuRI) map charts 52,569 interactions between 8,275 human proteins, as described in a study published in *Nature*.

Humans have about 20,000 protein-coding genes but scientists still know remarkably little about most of the proteins they encode. Fortunately, this information can be gleaned from interaction data thanks to the "guilt by association" principle, according to which two proteins that have similar interacting partners are likely involved in similar biological processes.

"We can use our human interactome map to predict protein function," says Roth, who is also Senior Scientist at the Sinai Health System's Lunenfeld-Tanenbaum Research Institute. "People can look up their favourite protein and get clues about its function from the proteins it interacts with."

The data are already revealing important insights such as new cellular roles for human proteins and what goes wrong at the molecular level to spur on disease.

In this vein, HuRI has already revealed new

functions for proteins involved in programmed cell death, release of cellular cargo and other processes.

And, by integrating protein interaction data with tissue-specific gene expression, the teams have been able to identify protein networks behind the development and maintenance of different tissues, revealing new therapeutic targets for diverse genetic diseases including cancer and potentially for infectious diseases as well.

Furthermore, using HuRI as a reference, they were also able to see how disease-causing protein variants bring about network rewiring to reveal molecular mechanisms behind those particular disorders.

"Genome sequencing can identify the variants carried by an individual that make them susceptible to disease, but it doesn't reveal how the disease is caused," says Mike Calderwood Ph.D., Scientific Director of the Center for Cancer Systems Biology (CCSB) at Dana-Farber Cancer Institute "Changes in the interactions of a protein is one possible mechanism of disease, and this map provides a starting point to study the impact of disease associated variants on protein-protein interactions."

The Toronto and Boston teams previously did two smaller studies mapping a total of ~14,000 protein interactions. Now HuRI has interrogated proteins encoded by nearly all human protein-coding genes and expanded the map four-fold.

To create HuRI, the researchers co-expressed in pairs almost all human proteins in yeast cells. When the two proteins interact, or bind one another, they form a molecular switch which boosts yeast cell growth—a sign that an interaction has occurred.

The team tested all possible pairwise combinations among 17,500 proteins for their ability to interact with each other in three separate versions of a yeast-based assay, each done in triplicate, amounting to a staggering three billion separate tests. The results yielded ~53,000 high-confidence binary interactions between more than 8,000 proteins, which were verified by other methods. The majority of interactions had never been detected before.

Although the largest map of its kind to date, the map remains incomplete, representing between 2-11 per cent of all human protein interactions. Roth said that one reason why many interactions were missed is probably because yeast cells lack certain human-specific molecular factors that are needed for proper protein function.

Despite these limitations, HuRI has more than tripled the number of known interactions between human proteins and will serve as an important resource for the research community. Already 15,000 people have visited the data web portal, which was built by Miles Mee, Mohamed Helmy, and Gary Bader (also in the Donnelly Centre), since HuRI was made available on bioRxiv, an open-source online publisher, in April 2019.

"We already had lots of people download the whole dataset and so I imagine we'll see the iteration of our previous paper, which has already been cited over 800 times and is less than a third of the size of HuRI," says Roth.

Provided by University of Toronto