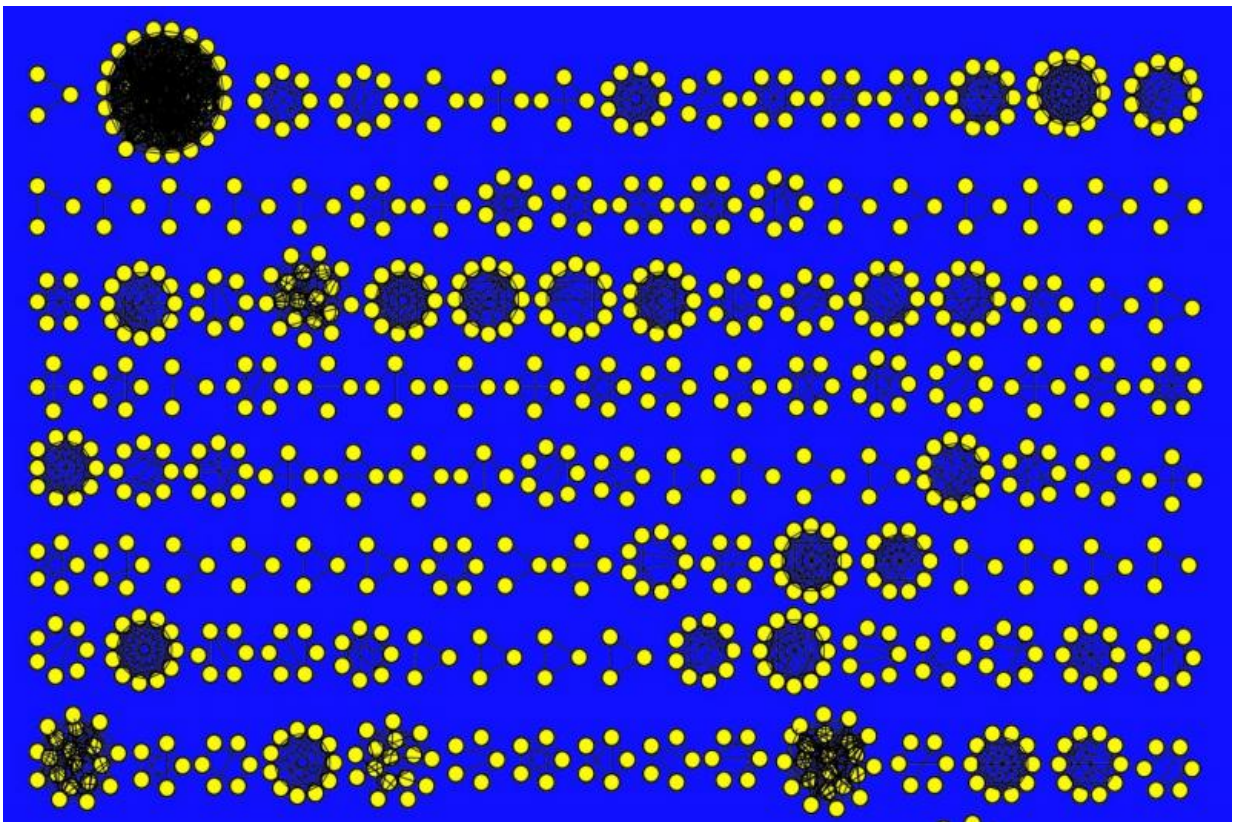# Scientists create world's largest protein map to reveal which proteins work together in a cell

September 7 2015

Scientists have uncovered tens of thousands of new protein interactions, accounting for about a quarter of all estimated protein contacts in a cell. Credit: Jovana Drinkjakovic

A multinational team of scientists have sifted through cells of vastly different organisms, from amoebae to worms to mice to humans, to reveal how proteins fit together to build different cells and bodies.

This tour de force of protein science, a result of a collaboration between seven research groups from three countries, led by Professor Andrew Emili from the University of Toronto's Donnelly Centre and Professor Edward Marcotte from the University of Texas at Austin, uncovered tens of thousands of new protein interactions, accounting for about a quarter of all estimated protein contacts in a cell.

When even a single one of these interactions is lost it can lead to disease, and the map is already helping scientists spot individual proteins that could be at the root of complex human disorders. The data will be available to researchers across the world through open access databases.

The study comes out in *Nature* on September 7.

While the sequencing of the human genome more than a decade ago was undoubtedly one of the greatest discoveries in biology, it was only the beginning of our in-depth understanding of how cells work. Genes are just blueprints and it is the genes' products, the proteins, that do much of the work in a cell.

Proteins work in teams by sticking to each other to carry out their jobs. Many proteins come together to form so called molecular machines that play key roles, such a building new proteins or recycling those no longer needed by literally grinding them into reusable parts. But for the vast majority of proteins, and there are tens of thousands of them in human cells, we still don't know what they do.

This is where Emili and Marcotte's map comes in. Using a state-of-the-art method developed by the groups, the researchers were able to fish

thousands of protein machineries out of cells and count individual proteins they are made of. They then built a network that, similar to social networks, offers clues into protein function based on which other proteins they hang out with. For example, a new and unstudied protein, whose role we don't yet know, is likely to be involved in fixing damage in a cell if it sticks to cell's known "handymen" proteins.

Today's landmark study gathered information on protein machineries from nine species that represent the tree of life: baker's yeast, amoeba, sea anemones, flies, worms, sea urchins, frogs, mice and humans. The new map expands the number of known protein associations over 10 fold, and gives insights into how they evolved over time.

"For me the highlight of the study is its sheer scale. We have tripled the number of protein interactions for every species. So across all the animals, we can now predict, with high confidence, more than 1 million protein interactions - a fundamentally 'big step' moving the goal posts forward in terms of protein interactions networks," says Emili, who is also Ontario Research Chair in Biomarkers in Disease Management and a professor in the Department of Molecular Genetics.

The researchers discovered that tens of thousands of protein associations remained unchanged since the first ancestral cell appeared, one billion years ago (!), preceding all of animal life on Earth.

"Protein assemblies in humans were often identical to those in other species. This not only reinforces what we already know about our common evolutionary ancestry, it also has practical implications, providing the ability to study the genetic basis for a wide variety of diseases and how they present in different species," says Marcotte.

The map is already proving useful in pinpointing possible causes of human disease. One example is a newly discovered molecular machine,

dubbed Commander, which consists of about a dozen individual proteins. Genes that encode some of Commander's components had previously been found to be mutated in people with intellectual disabilities but it was not clear how these proteins worked.

Because Commander is present in all animal cells, graduate student Fan Tu went on to disrupt its components in tadpoles, revealing abnormalities in the way brain cells are positioned during embryo development and providing a possible origin for a complex human condition.

"With tens of thousands of other new protein interactions, our map promises to open many more lines of research into links between proteins and disease, which we are keen to explore in depth over the coming years," concludes Dr. Emili.

Provided by University of Toronto