

DECEMBER 05, 2000

Untangling the Web of Yeast Protein Interactions

As researchers come to terms with the massive amount of data left in the wake of large-scale gene sequencing efforts, one of their first tasks is to begin assigning functions to myriad proteins produced by the genes. Howard Hughes Medical Institute (HHMI) researchers at the University of Washington have assembled the first map of protein interaction networks by analyzing published data on the interactions among thousands of yeast proteins. While the scientists caution that the map likely contains errors and omissions, they believe that their strategy represents a first step toward providing scientists with a reference guide to aid detailed exploration of the functions of yeast proteins.

The research team, led by HHMI investigator Stanley Fields at the University of Washington, reported in an article in the December 2000 issue of the journal *Nature Biotechnology* that it had analyzed more than 2,709 protein interactions, involving 2,039 different proteins, all of which had been published by the community of yeast researchers. Analyses done by Fields, Benno Schwikowski of the Institute for Systems Biology in Seattle and Peter Uetz, a former HHMI associate at the University of Washington, led to a surprise: the majority of protein interactions could be mapped into a single large network of 2,358 interactions that involved 1,548 proteins.

According to Fields, maps of protein interactions will initially be useful in guiding molecular and genetic studies that are needed to determine the function of proteins. "Mapping protein interactions can complement other analytical approaches to assigning function and characterizing proteins," he said. "At best, protein interaction data allow you to place an uncharacterized protein by associating it with other proteins in a metabolic pathway, a cellular structure, or a macromolecular complex. Once you have that information, you can carry out experiments to understand that protein based on those associations."

"Analysis of these networks allows assignment of potential function to uncharacterized proteins and the discovery of potential interactions within and across cellular processes and compartments," write Melanie Mayer of The Johns Hopkins University and Philip Hieter of the University of British Columbia in a *News and Views* article in *Nature Biotechnology*. "These connections represent a gold mine for formulating and experimentally testing specific hypotheses about gene function."

Fields said that he was initially skeptical about whether mapping these interactions would yield scientifically useful results. "A large number of interactions had been identified by researchers, but they had been displayed in tabular form or in lots of separate papers that looked at specific proteins," he said. "The question in my mind was whether a network view of the known interactions of yeast proteins would yield new information. The analysis and visualization of these interactions by Benno and Peter have now convinced me that it does."

The researchers tested the validity of their interaction map by assessing how reliable the map was in allowing them to predict the function of previously characterized proteins. They found that the network yielded a correct prediction for 72 percent of the 1,393 characterized proteins.

The scientists also discovered intriguing hints that proteins involved in specific functions, such as RNA processing, seemed to link to proteins involved in mitosis, chromatin and protein synthesis, as well as to proteins in expected functions such as RNA splicing, RNA turnover and RNA polymerase II transcription. Also, the network revealed some surprising clusters of interaction between proteins in distinct cellular compartments, such as the nucleus and cytoplasm.

While such associations are intriguing, Fields emphasized that the current network is likely to contain some interactions that are "false positives," since they might depend on results from only one methodology.

"This is not a real picture of what a cellular network is," he said. "For example, what looks like one giant network of proteins might be lots of smaller networks. This could happen if we have included a link that is false positive. Also, some interactions we've included might be mutually exclusivesuch that if one interaction occurs, it might preclude another. Finally, although we have this large network of 2,358 interactions, the real number of protein interactions in yeast is certainly much larger," he said.

"However, given these caveats, we were still surprised at how few proteins there were that weren't part of any pathway or machine or complex," said Fields.

While the network represents a beginning, said Fields, "it's clear that over the next few years, the yeast research community will continue to produce more protein interaction data, and we can use that information to build more complex networks." Likewise, parallel efforts to map protein interactions in the roundworm *C. elegans*, the fruit fly *Drosophila* and humans will yield more insight into understanding protein function.

And it's likely that future versions of the yeast protein interaction map will be aided by advances in computer graphics, which will allow the scientists to display the map in three dimensions. Fields predicts that future releases of the map will contain hyperlinks that direct users to descriptive data about each protein, making such networks even more useful as reference tools to guide research.

"Ultimately one would like to superimpose on these interaction maps cellular structures such as the nucleus, cytoplasm and plasma membrane, making the maps one component of a virtual cell that contains a variety of information about its molecules and structures," he said. "These virtual cells could prove to be useful tools to help biologists understand the processes, the proteins involved and how they are regulated."