# Secrets of human protein interactions unveiled by massive sequencing and coevolution

December 29 2016

Cells operate like incredibly well-synchronized orchestras of molecular interactions among proteins. This molecular network is essential not only to understanding how an organism functions, but also to determining the molecular mechanisms responsible for a multitude of diseases. In fact, it has been observed that protein interaction regions are preferentially mutated in tumours.

A study coordinated by Simone Marsili and David Juan, from Alfonso Valencia's team at the CNIO, published in the journal *Proceedings of the National Academy of Sciences (PNAS)*, demonstrates that it is possible to understand a significant number of interactions among human proteins from the evolution of their counterparts in simpler cells, such as bacteria.

According to Juan Rodríguez from the Structural Computational Biology Group at the CNIO and first author of the paper, "The complexity of human beings does not only result from the number of proteins that we have, but primarily from how they interact with each other. However, out of 200,000 protein-protein interactions estimated, only a few thousand have been characterised at the molecular level." It is very difficult to study the molecular properties of many important interactions without reliable structural information. Researchers are now exploring this "twilight zone" for the first time.

## Studying bacteria to understand human diseases

Although millions of years of evolution separate bacteria and humans, the CNIO team has utilized the information accumulated over thousands of bacterial sequences to predict interactions between proteins in humans. "We have used the protein coevolution phenomenon: Proteins that interact tend to experience coordinated evolutionary changes that maintain the interaction despite the accumulation of mutations over time," says David Juan. "We have demonstrated that we can use this phenomenon to detect molecular details of interactions in humans that we share with very distant species. What is most interesting is that this allows us to transfer information from bacteria in order to study interactions in humans that we knew almost nothing about," adds Simone Marsili.

These new results have important implications for future research. "A deeper understanding of these interactions opens the door to the modeling of three-dimensional structures that may help us to design drugs targeting important interactions in various types of cancer," explains David Juan.

"This knowledge can also improve our predictions of the effects of various mutations linked to tumour development," says Rodríguez.

## Data-based science

The laboratory of Alfonso Valencia, head of the Structural Biology and Biocomputing Programme, has been engaged with protein coevolution since the 1990s. This field has significantly advanced in recent years. "Thanks to the amount of biological data that is being generated today, we can use new computational methods that account for a greater number of factors," explains Valencia. According to the researchers, the

pace of innovation in massive experimental techniques is providing additional data, making it possible to design more complex statistical models that provide a more complete view of biological systems, something particularly important in multifactorial diseases, such as cancer.

**More information:** Juan Rodriguez-Rivas et al, Conservation of coevolving protein interfaces bridges prokaryote–eukaryote homologies in the twilight zone, *Proceedings of the National Academy of Sciences* (2016). DOI: 10.1073/pnas.1611861114

Provided by Centro Nacional de Investigaciones Oncológicas (CNIO)