

Using computational biology for the annotation of proteins

November 26 2012

Anotación Funcional



Image illustrating annotating proteins. Credit: UC3M

Research carried out at Universidad Carlos III of Madrid in collaboration with the Centro Nacional de Investigaciones Oncologicas employed computational techniques to improve the characterization of

proteins. The system they developed has allowed them to predict, for example, the relationship between two human proteins and telomeres, which led to their possible implication in cellular aging and the development of cancer; this awaits experimental verification.

Proteins are molecules that are formed by chains of [amino acids](#) and they play a fundamental role in all of life, given that they contain the coded information in genes; they, therefore, carry out numerous functions in an organism: immunological (antibodies), structural (they constitute the majority of cellular material), bioregulating (they form part of enzymes) and a long list of etceteras. In short, they regulate thousands of process that take place within all organisms, including inside the [human organism](#), and they frequently do so by means of relationships they establish with other cells. "Analyzing and using this network of interactions is a very interesting task due to the large number of associations that exist and to the multiple forms in which one [protein](#) can influence the function of others," explains Professor Beatriz García, of UC3M's Computer Science department. "In such a complex biological scenario, determining the functional associations through experiments is very costly, so we have tried to apply [computational tools](#) to predict these functions and so orient experimentation," she points out. Thus, the idea is to use techniques from the field of [Artificial Intelligence](#), specifically from the area of Machine Learning, to obtain useful results for Biology, as part of an emerging interdisciplinary field known as Biocomputing or [Computational Biology](#).

In this context, this line of research goes further in the annotation of the function of proteins, that is, in the determination of which protein or which group of proteins performs which task within an organism. In short, these scientists have dealt with two specific problems: the prediction of functional associations between pairs of proteins in the bacteria *Escherichia coli* and the extension of biological pathways in humans. In addition, they offer conclusions regarding the interpretation

of those predictions, which may help explain the function of the cellular processes that were studied. "In particular," states Beatriz García, "the predictions obtained regarding two human proteins stand out (E3 SUMO-protein ligase y E3 ubiquitin-protein ligase DTX1); these were previously related to the controlled degradation of certain proteins, and we can now propose a new function related to the stabilization of [telomeres](#) and, therefore, their possible implication in [cellular aging](#) and the development of cancer, which will require [experimental verification](#)."

For this study, part of which was recently published in the journal *PLOS ONE*, the researcher has received the award for the best doctoral dissertation in her field (Experimental Sciences and Technology) from the Real Academia de Doctores de España (Spanish Royal Academy of Doctors). The implications that this work holds for the scientific community are already being felt. In fact, the results of the first problem that the project analyzes have already been integrated into the predictions server EcID (*E.coli* Interaction Database) and they offer a reliability value for the predictions that improves the system's performance when finding functional associations among the proteins that appear in this database. Moreover, the second biological problem dealt with in the study opens a new area of research in Biocomputing, by extending already existing pathways. "The procedure it presents complements the only previously existing publication in the field, extending the pathways with many more proteins and exploring a greater surface of the network of interactions," comments the researcher. In addition, it could be applied to many more problems of functional annotation in Biology and other fields in which there is relevant information with multiple relationships.

In any case, much work remains to be done in the area of Biocomputation. "There are still so many unresolved biological problems that need computational solutions," assures Beatriz García, who

highlights the relevance of this field, which is growing with the advances in new technologies; yet many computational challenges remain, such as the analysis of the new generation of sequencing. "This is an area that needs more trained professionals who can integrate Biology and Computer Science, in order to improve our knowledge of our organism at the molecular level and, finally, to facilitate the treatment of diseases," she concludes.

More information: Inference of Functional Relations in Predicted Protein Networks with a Machine Learning Approach, Beatriz García-Jiménez, David Juan, Iakes Ezkurdia, Eduardo Andres-León, Alfonso Valencia, *PLOS ONE* 5(4): e9969. PLOS ONE 2010; [doi: 10.1371/journal.pone.0009969](https://doi.org/10.1371/journal.pone.0009969)

Provided by Carlos III University of Madrid

Citation: Using computational biology for the annotation of proteins (2012, November 26) retrieved 25 April 2024 from <https://phys.org/news/2012-11-biology-annotation-proteins.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>
--