

New website to help translate genetic data into medical therapies

June 13 2009

Princeton researchers have created a Rosetta Stone for the human body, a website that offers clues to the role DNA plays in aging and disease by helping scientists make sense of the vast jumble of information emerging from genetics research.

By mashing up genetic data from disparate sources and interpreting it with the help of computer algorithms informed by biological principles, the online system allows scientists to predict which genes might be involved in ailments such as Alzheimer's disease, diabetes and cancer.

"The scientific community has produced millions of points of [genetic data](#) in recent years, but has not achieved an equivalent understanding of how genes work," said Olga Troyanskaya, the Princeton professor who led the project. "We need to translate this into knowledge about disease."

Reflecting Troyanskaya's joint appointments as an assistant professor in the Department of Computer Science and the Lewis-Sigler Institute for Integrative Genomics, the new website exists at the nexus of computers and genomics, the field of biology concerned with mapping organisms' entire DNA and understanding how genes interact to keep an organism healthy or cause disease.

"Olga has now emerged as a world leader in analyzing and displaying vast amounts of functional data so that the ordinary biologist can understand them," said David Botstein, the Anthony B. Evnin Professor of Genomics and director of the Lewis-Sigler Institute.

In conjunction with launching the new site -- which was developed by Curtis Huttenhower, a postdoctoral researcher in Troyanskaya's lab -- the team's paper on its methodology, titled "Exploring the [Human Genome With Functional Maps](#)," was published in the May issue of the journal *Genome Research*.

The site is based on the principle of "functional mapping." The term is shorthand for mapping out the tangled web of relationships among genes, based on how they work together in cellular function. A single gene, for example, might help a cell become heart or brain tissue, but a cell's overall function emerges from the interactions of many genes.

Understanding these functional relationships is key to developing new medical treatments, since most medications target proteins -- the primary product of genes. Proteins are complex molecules that serve as cogs in the cellular machinery or, in the case of disease, wrenches in the works.

Genomics researchers seek to understand which genes and proteins are involved in certain aspects of cell function. Is a protein part of the mechanism that produces energy for the cell? Does it work in concert with other genes to control aging? Does it help control the metabolic rhythms that serve as the basis of humans' biological clocks?

Working out how genes keep cells running normally helps scientists understand what goes wrong in the case of a harmful genetic mutation. Discovering a link between a gene and a disease can tell researchers what cellular processes are involved in the disease, which in turn fingers other genes involved in those processes as potential culprits.

But discerning these connections is no easy feat. Discoveries of genes resemble early discoveries of Egyptian hieroglyphs: Finding a new one doesn't mean researchers understand its purpose or how it fits into the larger system.

While Egyptologists struggled to decode the meaning of around 2,000 hieroglyphs, genomics researchers are faced with an estimated 20,000 to 25,000 human genes that could potentially interact with each other in 300 million different ways.

With so many genes and so many possible avenues of inquiry, predicting which genes and relationships are important in certain diseases, and therefore worthwhile to study, presents an enormous challenge. It involves a lot of guesswork.

This is where computers come in handy. The computer program created by Troyanskaya and the other computational biologists working on the project sorts through 350 sets of genome data from thousands of separate experiments.

The program relies on artificial intelligence algorithms, similar to those used by government intelligence agencies to sort through the data collected as part of anti-terrorism programs and by online commerce websites, such as Amazon and Netflix, to recommend products to customers.

Dubbed the Human Experimental/Functional Mapper, or HEFaIMp, the site focuses on discerning connections among genes, biological processes and diseases to help scientists determine which relationships are most important.

Entering "breast cancer," for instance, returns a list of all the genes in the site's database ordered by the probability that they are involved in the development of the disease. Three genes at the top of the list -- BRCA1, BRCA2 and TP53 -- are known to play an important role in the development of breast cancers, but other genes high on the list also could be involved. The site allows researchers to explore how these genes work together and the likely reasons they play a role in breast cancer.

"Knowing which genes are most likely to be involved helps researchers choose where to focus," Troyanskaya said. "The program determines the significance between a gene and a disease based on a rigorous analysis of published data."

"This is a magnifying glass," she said, "that shows you what is trustworthy and what is relevant."

Troyanskaya anticipates that molecular biologists will begin using the site following publication of the paper. Hilary Collier, an assistant professor of molecular biology at Princeton who co-wrote the paper with Troyanskaya, used the site to link [genes](#) to an important cellular process, known as autophagy, by which nutrient-starved cells digest parts of themselves to ensure survival. The results of the laboratory tests were published in the paper.

Members of Collier's research group continue to use the site to understand the results of their laboratory experiments and to provide clues to new avenues of research.

"In the past, everyone did their own experiments and came to their own conclusions," she said. "It was rare that anyone actually compared results, in part because it was overwhelming. There was always this sense that if someone pulled all of this information together it would be valuable. The new site does an intelligent job of mining a lot of data and putting it into an intelligible form."

Source: Princeton University ([news](#) : [web](#))

Citation: New website to help translate genetic data into medical therapies (2009, June 13) retrieved 17 April 2024 from

<https://medicalxpress.com/news/2009-06-website-genetic-medical-therapies.html>

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.