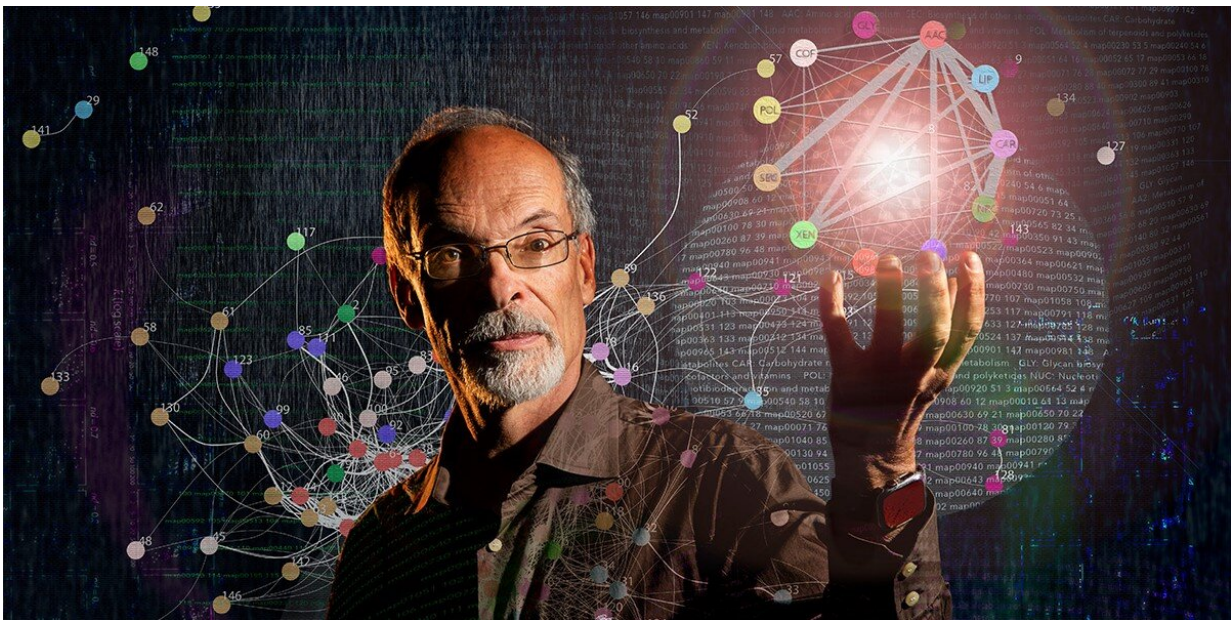


# Protein 'big bang' reveals molecular makeup for medicine and bioengineering

June 30 2021, by Lauren Quinn



Gustavo Caetano-Anolles surrounded by depiction of molecular networks.  
Credit: Fred Zwicky.

Proteins have been quietly taking over our lives since the COVID-19 pandemic began. We've been living at the whim of the virus's so-called "spike" protein, which has mutated dozens of times to create increasingly deadly variants. But the truth is, we have always been ruled by proteins. At the cellular level, they're responsible for pretty much everything.

Proteins are so fundamental that DNA—the genetic material that makes each of us unique—is essentially just a long sequence of [protein](#) blueprints. That's true for animals, plants, fungi, bacteria, archaea, and even viruses. And just as those [groups of organisms](#) evolve and change over time, so too do proteins and their component parts.

A new study from University of Illinois researchers, published in *Scientific Reports*, maps the [evolutionary history](#) and interrelationships of protein domains, the subunits of protein molecules, over 3.8 billion years.

"Knowing how and why domains combine in proteins during evolution could help scientists understand and engineer the activity of proteins for medicine and bioengineering applications. For example, these insights could guide disease management, such as making better vaccines from the spike protein of COVID-19 viruses," says Gustavo Caetano-Anollés, professor in the Department of Crop Sciences, affiliate of the Carl R. Woese Institute for Genomic Biology at Illinois, and senior author on the paper.

Caetano-Anollés has studied the evolution of COVID mutations since the early stages of the pandemic, but that timeline represents a vanishingly tiny fraction of what he and doctoral student Fayez Aziz took on in their current study.

The researchers compiled sequences and structures of millions of [protein sequences](#) encoded in hundreds of genomes across all taxonomic groups, including higher organisms and microbes. They focused not on whole proteins, but instead on structural domains.

"Most proteins are made of more than one domain. These are compact structural units, or modules, that harbor specialized functions," Caetano-Anollés says. "More importantly, they are the units of evolution."

After sorting proteins into domains to build evolutionary trees, they set to work building a [network](#) to understand how domains have developed and been shared across proteins throughout billions of years of evolution.

"We built a time series of networks that describe how domains have accumulated and how proteins have rearranged their domains through evolution. This is the first time such a network of 'domain organization' has been studied as an evolutionary chronology," Fayez Aziz says. "Our survey revealed there is a vast evolving network describing how domains combine with each other in proteins."

Each link of the network represents a moment when a particular domain was recruited into a protein, typically to perform a new function.

"This fact alone strongly suggests domain recruitment is a powerful force in nature," Fayez Aziz says. The chronology also revealed which domains contributed important protein functions. For example, the researchers were able to trace the origins of domains responsible for environmental sensing as well as secondary metabolites, or toxins used in bacterial and plant defenses.

The analysis showed domains started to combine early in protein evolution, but there were also periods of explosive network growth. For example, the researchers describe a "big bang" of [domain](#) combinations 1.5 billion years ago, coinciding with the rise of multicellular organisms and eukaryotes, organisms with membrane-bound nuclei that include humans.

The existence of biological big bangs is not new. Caetano-Anollés' team previously reported the massive and early origin of metabolism, and they recently found it again when tracking the history of metabolic networks.

The historical record of a big bang describing the evolutionary

patchwork of proteins provides new tools to understand protein makeup.

"This could help identify, for example, why structural variations and genomic recombinations occur often in SARS-CoV-2," Caetano-Anollés says.

He adds that this new way of understanding proteins could help prevent pandemics by dissecting how virus diseases originate. It could also help mitigate disease by improving vaccine design when outbreaks occur.

**More information:** M. Fayez Aziz et al, Evolution of networks of protein domain organization, *Scientific Reports* (2021). [DOI: 10.1038/s41598-021-90498-8](https://doi.org/10.1038/s41598-021-90498-8)

Provided by University of Illinois at Urbana-Champaign

Citation: Protein 'big bang' reveals molecular makeup for medicine and bioengineering (2021, June 30) retrieved 29 January 2023 from <https://phys.org/news/2021-06-protein-big-reveals-molecular-makeup.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.