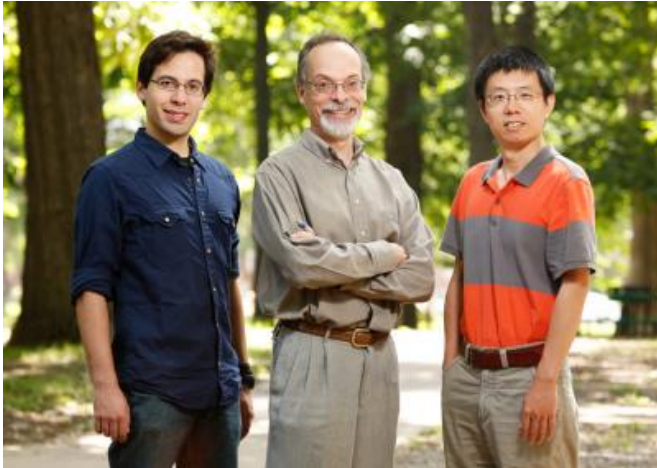


Study offers insight into the origin of the genetic code

26 August 2013



Graduate student Derek Caetano-Anollés, crop sciences and Institute for Genomic Biology professor Gustavo Caetano-Anollés and senior bioinformatician Minglei Wang report that the emergence of the genetic code corresponds to the advent of protein flexibility. Credit: L. Brian Stauffer

An analysis of enzymes that load amino acids onto transfer RNAs—an operation at the heart of protein translation—offers new insights into the evolutionary origins of the modern genetic code, researchers report. Their findings appear in the journal *PLOS ONE*.

The researchers focused on aminoacyl tRNA synthetases, enzymes that "read" the genetic information embedded in transfer RNA molecules and attach the appropriate amino acids to those tRNAs. Once a tRNA is charged with its amino acid, it carries it to the [ribosome](#), a cellular "workbench" on which proteins are assembled, one amino acid at a time.

Synthetases charge the amino acids with high-energy [chemical bonds](#) that speed the later formation of new peptide (protein) bonds.

Synthetases also have powerful editing capabilities; if the wrong amino acid is added to a tRNA, the enzyme quickly dissolves the bond.

"Synthetases are key interpreters and arbitrators of how nucleic-acid information translates into amino-acid information," said Gustavo Caetano-Anollés, a University of Illinois professor of [crop sciences](#) and of bioinformatics. Caetano-Anollés, who led the research, also is a professor in the U. of I. Institute for Genomic Biology. "Their editing capabilities are about 100-fold more rigorous than the proofreading and recognition that occurs in the ribosome. Consequently, synthetases are responsible for establishing the rules of the genetic code."

The researchers used an approach developed in the Caetano-Anollés lab to determine the relative ages of different protein regions, called domains. Protein domains are the gears, springs and motors that work together to keep the [protein machinery](#) running.

Caetano-Anollés and his colleagues have spent years elucidating the evolution of protein and RNA domains, determining their relative ages by analyzing their utilization in organisms from every branch of the tree of life. The researchers make a simple assumption: Domains that appear in only a few organisms or groups of organisms are likely younger than domains that are more widely employed. The most universally utilized domains—those that appear in organisms from every branch of the tree of life—are likely the most ancient.

The researchers used their census of protein domains to establish the relative ages of the domains that make up the synthetases. They found that those domains that load amino acids onto the tRNAs (and edit them when mistakes are made) are more ancient than the domains that recognize the region on the tRNA, called an anticodon, that tells the synthetase which amino acid that tRNA

should carry.

"Remarkably, we also found that the most ancient domains of the synthetases were structurally analogous to modern enzymes that are involved in non-ribosomal protein synthesis, and to other enzymes that are capable of making dipeptides," Caetano-Anollés said.

The researchers hypothesize that ancient protein synthesis involved enzymes that looked a lot like today's synthetases, perhaps working in conjunction with ancient tRNAs.

Researchers have known for decades that rudimentary protein synthesis can occur without the involvement of the ribosome, Caetano-Anollés said. But few if any have looked to the enzymes that catalyze these reactions for evidence of the [evolutionary origins](#) of protein synthesis.

Alerted to the potential importance of dipeptide formation in early protein synthesis, the researchers next looked for patterns of frequently used dipeptides in the sequences of modern proteins. They focused only on proteins for which scientists have collected the most complete and accurate structural information.

"The analysis revealed an astonishing fact," Caetano-Anollés said. "The most ancient [protein domains](#) were enriched in dipeptides with [amino acids](#) encoded by the most ancient synthetases. And these ancient dipeptides were present in rigid regions of the proteins."

The domains that appeared after the emergence of the genetic code (which Caetano-Anollés ties to the emergence of the tRNA anticodon) "were enriched in dipeptides that were present in highly flexible regions," he said.

Thus, genetics is associated with protein flexibility, he said.

"Our study offers an explanation for why there is a [genetic code](#)," Caetano-Anollés said. Genetics allowed proteins "to become flexible, thereby gaining a world of new molecular functions."

More information: The paper, "Structural Phylogenomics Retrodicts the Origin of the Genetic Code and Uncovers the Impact of Protein Flexibility," is available online.

www.plosone.org/article/info%3Djournal.pone.0072225

Provided by University of Illinois at Urbana-Champaign

APA citation: Study offers insight into the origin of the genetic code (2013, August 26) retrieved 17 June 2019 from <https://phys.org/news/2013-08-insight-genetic-code.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.