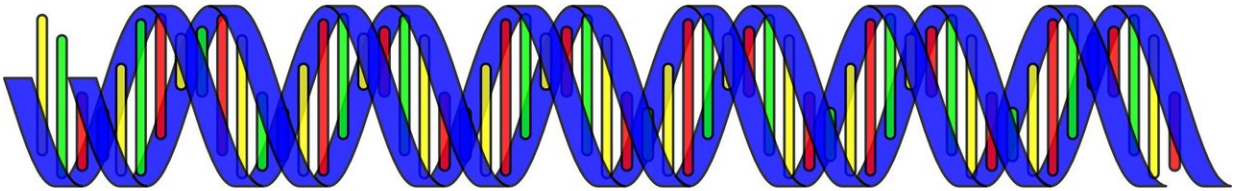


Genes that evolve from scratch expand protein diversity

March 11 2019, by Matt Wood



Credit: CC0 Public Domain

One of the most important questions in biology is how rapidly new proteins evolve in organisms. Proteins are the building blocks that carry out the basic functions of life. As the genes that produce them change, the proteins change as well, introducing new functionality or traits that can eventually lead to the evolution of new species.

A new study published in *Nature Ecology and Evolution* led by scientists

from the University of Chicago challenges one of the classic assumptions about how new proteins evolve. The research shows that random, noncoding sections of DNA can quickly evolve to produce new proteins. These de novo, or "from scratch," genes provide a new, unexplored way that proteins evolve and contribute to biodiversity.

"Using a big genome comparison, we show that noncoding sequences can evolve into completely novel proteins. That's a huge discovery," said Manyuan Long, Ph.D., the Edna K. Papazian Distinguished Service Professor of Ecology and Evolution at UChicago and senior author of the new study.

A third way for genes to evolve

For decades, scientists believed that there were only two ways new genes evolved: duplication and divergence or recombination. During the normal process of replication and repair, a section of DNA gets copied and creates a duplicate version of the gene. Then, one of these copies may acquire mutations that change its functionality enough that it diverges and becomes a distinct new gene. With recombination, pieces of genetic material are reshuffled to create new combinations and new genes. However, these two methods only account for a relatively small number of proteins, given the total number of possible combinations of amino acids that comprise them.

Scientists have long wondered about a third mechanism, where de novo genes could evolve from scratch. All organisms have long stretches of genetic material that do not encode proteins, sometimes up to 97 percent of the total genome. Is it possible for these noncoding sections to acquire mutations that suddenly make them functional?

This has been difficult to study because it requires high-quality reference genomes from several closely related species that show both

the ancestral, noncoding sequences and subsequent new genes that evolved from them. Without this clear, visible line of [evolution](#), there's no way to prove it's truly a de novo gene. The supposed new genes reported previously could just be an "orphaned gene" that diverged or transferred from unrelated organisms at some point, then all traces of its predecessors disappeared.

To overcome these challenges, Long's team took advantage of 13 new genomes sequenced and annotated recently from 11 closely-related species of rice plants, including *Oryza sativa*, the most common food crop. He worked with groups headed by Prof. Rod Wing at the University of Arizona. Prof. Yidan Ouyang from Huazhong Agricultural University, China, also led a team that cultivated their own rice plants in Hainan, a tropical island off the southern coast of China, and harvested them for proteomics sampling.

After analyzing the genomes of these plants, they detected at least 175 de novo genes. Further mass spectrometry analysis of [protein](#) activity was conducted by another group led by Prof. Siqi Liu at BGI-Shenzhen, a genome sequencing center located in Shenzhen, Guangdong, China. They found evidence that 57 percent of these genes actually translated into new proteins, including more than 300 new peptides.

With this first, large dataset of authentic de novo genes, Long's team detected a pattern in their evolution. It began with the early evolution of expression, followed by subsequent mutation into protein coding potentials for almost all de novo genes.

"This makes sense given the widely observed expression of intergenic regions in various organisms," said Li Zhang, a postdoctoral researcher at UChicago and lead author of the article.

Long says that the *Oryza* plants are good genomes to search for de novo

genes because they are relatively young—you can still see evidence of evolution in their existing genomes.

"The 11 species diverged from each other only about three to four million years ago, so they are all young species," he said. "For that reason, when we sequence the genomes, all the sequences are highly similar. They haven't accumulated multiple generations of changes, so all the previous non-coding sections are still there."

Long and his team next want to study the new proteins to further understand their function and evolution and see if there is something unique about their structure. If de novo genes open up an unexplored path for evolution, they could reveal mechanisms for creating new and improved cellular functions. For instance, the researchers detected evidence of natural selection acting to fix insertions and deletions in the genome to generate new protein sequences, and the sequence's evolution toward improved functions.

"The new proteins may make certain functions better, or help regulate the [genes](#) better," he said. "Each step of the way, they can bring some kind of benefit to the organism until it gradually becomes fixed in the [genome](#)."

More information: Rapid evolution of protein diversity by de novo origination in *Oryza*, *Nature Ecology and Evolution* (2019). [DOI: 10.1038/s41559-019-0822-5](https://doi.org/10.1038/s41559-019-0822-5), www.nature.com/articles/s41559-019-0822-5

Provided by University of Chicago Medical Center

Citation: Genes that evolve from scratch expand protein diversity (2019, March 11) retrieved 7

May 2024 from <https://phys.org/news/2019-03-genes-evolve-protein-diversity.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.