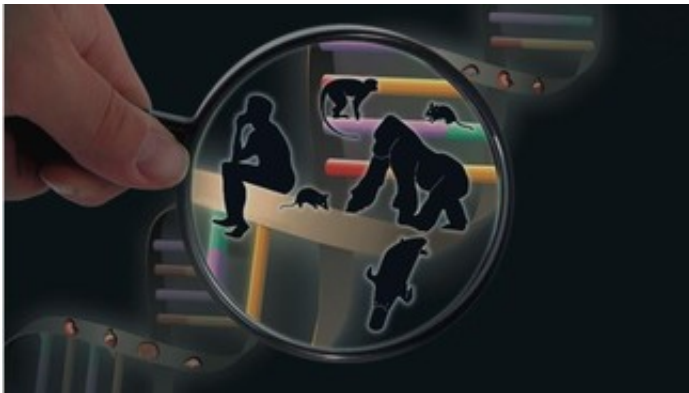


Scientists shed some light on biological "dark matter"

January 20 2014



Credit: Unil

Biologists have studied the functionality of a poorly understood category of genes, which produce long non-coding RNA molecules rather than proteins. Some of these genes have been conserved throughout evolution, and are present in 11 species ranging from man to frog. The research was lead at the University of Lausanne, in partnership with EPFL and the Swiss Institute of Bioinformatics (SIB -SIB). It has been published today in *Nature*.

The "classical" role for a gene is to produce proteins, which are essential for the functioning of cells. However, our genomes also encode genes that produce long non-coding RNAs, whose functions are more mysterious. Yet, since four or five years ago we know that thousands of

these still poorly understood genes are present in the human and mouse genomes. How and in which organs are they activated? Is this biological "dark matter" much ado about nothing or is there something more interesting to it?

A team led by Professor Henrik Kaessmann at UNIL's Center for Integrative Genomics (CIG) compiled an authentic catalog of long non-coding RNAs in eleven [species](#). By adopting an evolutionary approach, they discovered that about 2500 long non-coding RNAs first appeared at least 90 million years ago in the common ancestor of most [placental mammals](#). From a functional point of view, these "ancient" genes turned out to be particularly interesting.

First author of the *Nature* article, Anamaria Necșulea, scientist at the Laboratory of developmental genomics of EPFL, expanded the scope of the investigation on these long non-coding RNAs to six primate species (man, macaque, chimpanzee, bonobo, gorilla and orangutan), and to mouse, opossum (a marsupial mammal), platypus (a monotreme mammal that lays eggs and nurses its young with milk), as well as to an "external group" composed by a bird (chicken) and an amphibian (frog). The common ancestor of all these species goes back over 350 million years.

Genes maintained during evolution

The biologists used the CIG genomics platform and the Vital-IT computing center at the Swiss Institute of Bioinformatics to identify long non-coding RNAs in several major organs of the 11 species under scrutiny. "Thanks to bioinformatics, we discovered RNA sequences produced from genome locations where no genes had been previously mapped," she said. "We then analyzed these genes to find out whether or not they encoded proteins. Thus, we could identify between 3000 and 15000 long non-coding RNA genes, depending on the species."

In the second phase of the research, a comparison among the different species allowed the scientists to pinpoint the emergence of these genes in the evolutionary history. While 11,000 long non-coding RNAs are shared by all primates, 2,500 go back to an ancestor common to man and mouse, about 90 million years ago. Only a hundred genes of this kind stem from an ancestor common to all eleven species considered, including birds and amphibians. "One of our main findings is that the activity of these non-coding genes is controlled by the same transcription factors that regulate protein-coding gene activity. Even more strikingly, we found that the 2,500 oldest long noncoding RNA genes are regulated by factors that are important for embryonic development. This suggests that, among the 2500 long non-coding RNAs conserved during the evolution of placental mammals, a large percentage may function specifically in embryonic development."

New network of interactions

The third phase of the research allowed the scientists to highlight a network of interactions (specifically, co-expression interactions, that is: genes are activated in the same organs or cell types) involving both long non-coding RNAs and protein-coding genes. For instance, they found that some non-coding genes are strongly associated to protein coding genes involved in brain function or in spermatogenesis, which suggests similar functions for these long non-coding RNA genes.

In the case of the H19X gene - one of the most ancient long noncoding RNA genes identified in this study - its association to the placental mammals' H19 gene (which was the first long non-coding RNA identified years ago) helped to uncover its functioning: "The H19 prevents the placenta from excessively growing inside the mother's womb," said Anamaria Necsulea. "We can assume that H19X also contributes to this function. We now plan to disable this gene in mice to test its functions in the placenta."

Among the subcategories of RNA producing genes, are these long RNA genes more useful than it originally seemed? By tracking them in 11 different species, this new study of unprecedented scale suggests that some of our genomes' "dark matter" may play a role in the development and functioning of the most vital organs of our bodies. Future experimental studies will further clarify the role of these [genes](#) that have just revealed their first secrets to us.

More information: The evolution of lncRNA repertoires and expression patterns in tetrapods, [DOI: 10.1038/nature12943](https://doi.org/10.1038/nature12943)

Provided by Ecole Polytechnique Federale de Lausanne

Citation: Scientists shed some light on biological "dark matter" (2014, January 20) retrieved 23 April 2024 from <https://phys.org/news/2014-01-scientists-biological-dark.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.