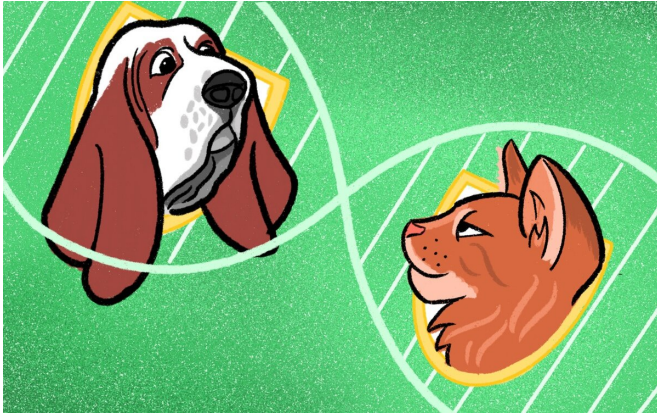


Scientists create program that finds synteny blocks in different animals

23 June 2020, by Itmo University



Credit: Dmitry Lisovsky, ITMO.NEWS

Modern genetics implies working with immense amounts of data which cannot be processed without the help of complex mathematical algorithms. For this reason, the task of developing special processing programs is no less important for bioinformatics specialists than that of genomic sequencing of specific animals. An international team of scientists that included researchers from ITMO University developed a software tool that makes it possible to quickly and efficiently find similar parts in the genomes of different animals, which is essential for understanding how closely related two species are, and how far they have evolved from their common ancestor. The research was published in *GigaScience*.

There are millions of biological species on planet Earth, and this diversity is laid down on the genetic level. Animals' anatomy, size, color patterns and habits are defined by their [genes](#). Then again, the diversity of genes themselves is not that great: by today, scientists have only identified about over 20,000. Therefore, species are different in not only the sets of genes they have but also in how their genes are arranged. In the language of

comparative genomics, this is called synteny, i.e. the arrangement of genes and regulatory elements.

"Let's take a gorilla and a chimpanzee as an example," says Ksenia Krasheninnikova, a researcher and engineer at ITMO University. "These two species have the same set of genes, but their regulatory elements and [genome](#) mutations create slightly different orders which results in differences between these primates."

Therefore, for the purposes of understanding how close two species are from the evolutionary standpoint, scientists need to know not just their genes but also how they are arranged in a chromosome, and how many common genome fragments, or synteny blocks, as geneticists call them, there are. Then again, looking for them manually is impossible: the amount of data is just too big. Genomes of mammals consist of millions and billions of base pairs, which makes processing without big data technologies next to impossible. For this reason, scientists create programs of their own that make it possible to solve this new category of tasks which has emerged in the course of the development of this science. And this is what the research team that included scientists from ITMO's Laboratory of Genomic Diversity did.

The new software tool was named halSynteny. According to its authors, it can search for synteny blocks better and faster than other programs developed for this purpose. What's more, halSynteny works with data in two standard and well-documented formats.

"Our goal was to create an algorithm that could be easily applied to accessible data," says Ksenia, who is the first author of this research. "Some of the approaches to the identification of synteny sequences are based on annotating genes in advance; our method is different. We don't use any additional annotation. We use the alignment

method, when different parts of one genome are aligned by their degree of similarity with parts of another genome. This way, we can identify homogeneous parts, parts that are of the same origin."

The program makes it possible to speed up the computations by over two times in comparison with SatsumaSynteny2, another popular tool. Such high efficiency was attained by implementing a mathematically effective algorithm using C++.

The proposed method and [software tool](#) were tested by comparing cat and dog genomes.

Ksenia says, "We showed that large fragments of cat chromosomes and some fragments of dog chromosomes unite in synteny blocks, which means that they've evolved from similar chromosomes of a common ancestor. And this can be used as a basis for making conclusions about their evolutionary process. Previous research in the field of 'wet' biology demonstrated that cats' genome changed less from the genome of their common ancestor in comparison with that of dogs. This can be seen in comparison with other species that are not part of the carnivora order. The results that we got confirm these conclusions and make them more accurate. This means that in some specific part, the genome of a cat and the [species](#) taken for comparison is similar, and in dogs, it is rearranged."

In future, this algorithm will be used in other research in the field of comparative genomics that takes place at ITMO University.

More information: Ksenia Krasheninnikova et al. halSynteny: a fast, easy-to-use conserved synteny block construction method for multiple whole-genome alignments, *GigaScience* (2020). [DOI: 10.1093/gigascience/giaa047](https://doi.org/10.1093/gigascience/giaa047)

Provided by ITMO University

APA citation: Scientists create program that finds synteny blocks in different animals (2020, June 23) retrieved 23 October 2020 from <https://phys.org/news/2020-06-scientists-synteny-blocks-animals.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.