

Personalised medicine will employ computer algorithms

June 16 2016



Artist's representation of MetaFast work process. Credit: Author Alla Tyakht

Scientists from ITMO University, the Federal Research and Clinical Centre of Physical-Chemical Medicine and MIPT have developed a software program that quickly compares sets of DNA of microorganisms living in different environments. Using the algorithm to compare the microflora of a healthy person with the microflora of a patient, specialists could detect previously unknown pathogens and their strains, which can aid the development of personalized medicine. The results of the study have been published in *Bioinformatics*.

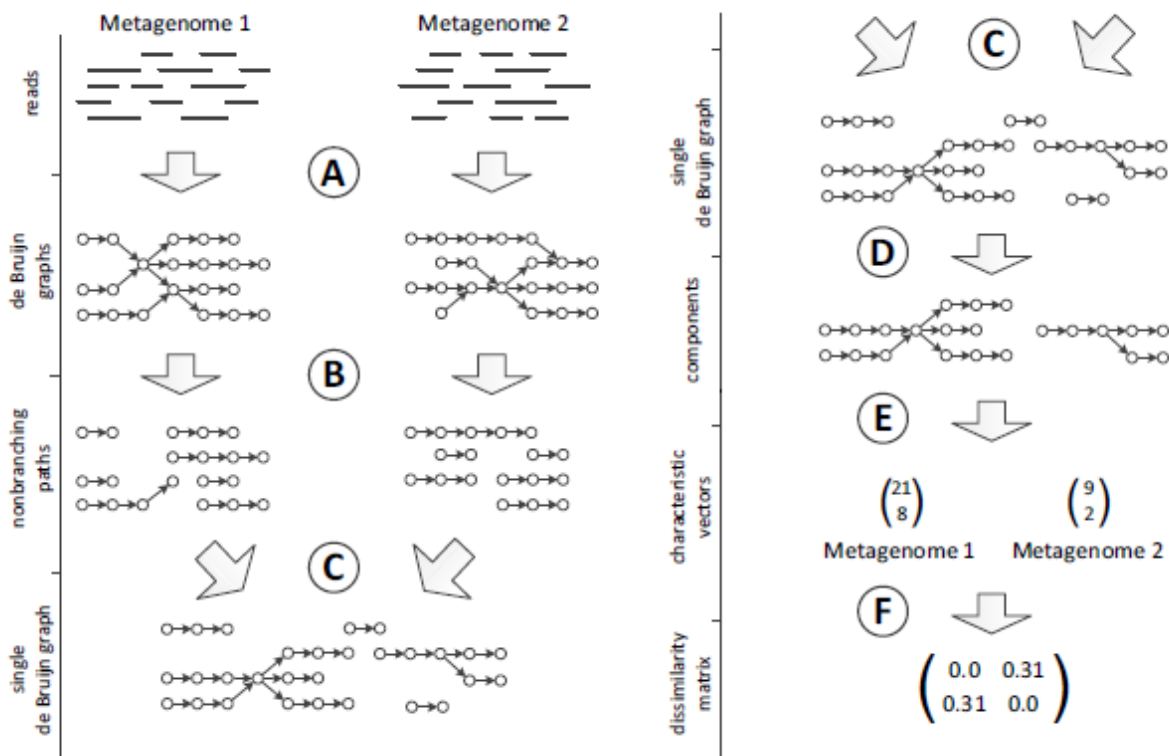
The genome is a specific sequence of genes according to which an individual develops. However, there is another gene sequence called the metagenome—the total DNA content of the many microorganisms that inhabit the same environment, including bacteria, fungi and viruses. The metagenome is often indicative of diseases or predispositions to diseases. Studying microbiota, i.e. the full range of microorganisms inhabiting various parts of the human body, is critical in metagenomic research.

The software developed by the scientists is called [MetaFast](#), and it can conduct a rapid comparative analysis of large numbers of metagenomes. "In studying the intestinal microflora of patients, we may be able to detect microorganisms associated with a particular disease, such as diabetes, or a predisposition to the disease. This forms a basis for applying personalized medicine techniques and developing [new drugs](#). Using the results obtained with the software, biologists will be able to draw conclusions on how to further develop their research, because the algorithm enables them to study environments that we currently know nothing about," says Vladimir Ulyantsev, lead developer of the algorithm and researcher at the Computer Technologies Laboratory at ITMO University.

One of the key benefits of the program is that it works successfully with environments in which the genetic contents have not yet been studied. "The approach allows us to do two things—find all the possible gene

sequences, even if they were previously unknown (the program collects them from fragments of genomic reads), and at the same time identify metagenomic patterns that distinguish one patient from another, e.g. people with and without a disease," says Dmitry Alexeev, the leader of the project and head of MIPT's Laboratory of Complex Biological Systems.

This means that the program can be used to conduct an untargeted express analysis of markers indicating certain diseases. Then, by using targeted methods such as PCR (a technique to make multiple copies of DNA fragments), the results can be verified and adjusted. According to the researchers, the program could greatly reduce the time needed to develop new drugs.



A flowchart of the MetaFast algorithm. Credit: ITMO University

Microorganisms that do not reproduce in vitro, such as viruses, give very abstract results in tests and it is not possible to collect their DNA. However, the new program can detect even these microorganisms. "In the microbiota of the skin alone, 90 percent of the organisms are unknown," continues Dmitry Alexeev. "Our approach enables us to work with completely unknown material and still obtain results. The program has been tested in a wide variety of environments, including those with a high number of viruses. The program can even locate and collect single DNA strands."

MetaFast is not limited to detecting pathogens. For example, the program can also be used to compare distinct peoples in closed populations with people living in cities to identify bacterial strains that are extremely useful to humans, but have possibly been lost in the process of urbanization. Antibiotics, preservatives, colorants and supermarket food have pushed many useful bacteria out of our microflora, which could still be present in closed populations such as American Indians or people in Russian villages.

MetaFast has proven to be highly effective in studying rare and undiscovered metagenomes. As a part of the study, the scientists analysed the metagenome of several of the world's largest lakes. Without any information about the samples of microbiota from the lakes, the program found genetic similarities between samples that were close in terms of their chemical composition.

The researchers also used the [new algorithm](#) to study the inhabitants of the New York subway system, demonstrating the effectiveness of the algorithm for analysing such complex systems. Most of the DNA

collected using MetaFast belonged to known bacteria. This confirms previous theories stating that the subway is safe for humans, and the microbes that live there suppress any flora that could be dangerous to people.

A vast amount of experimental data has already been gathered worldwide on various metagenomes. As the cost of extracting DNA is decreasing and the sensitivity of equipment is increasing, the volume of data is growing exponentially. Despite this, most of the studies have not been fully completed. The reason lies in the limitations of the current technology. On the one hand, scientists can collect a partial metagenome, but piecing together the whole puzzle takes an enormous amount of time. On the other hand, they can compare individual fragments of the genome with existing DNA references, but there are very limited numbers of bacteria, and virtually no viruses.

The new algorithm not only combines the advantages of both of these approaches, but also enables high-speed data processing. The program saves RAM because it partially collects and partially compares genomes, but does not go into an in-depth collection analysis.

More information: Vladimir I. Ulyantsev et al, MetaFast: fast reference-free graph-based comparison of shotgun metagenomic data, *Bioinformatics* (2016). [DOI: 10.1093/bioinformatics/btw312](https://doi.org/10.1093/bioinformatics/btw312)

Provided by ITMO University

Citation: Personalised medicine will employ computer algorithms (2016, June 16) retrieved 25 April 2024 from <https://phys.org/news/2016-06-personalised-medicine-algorithms.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.