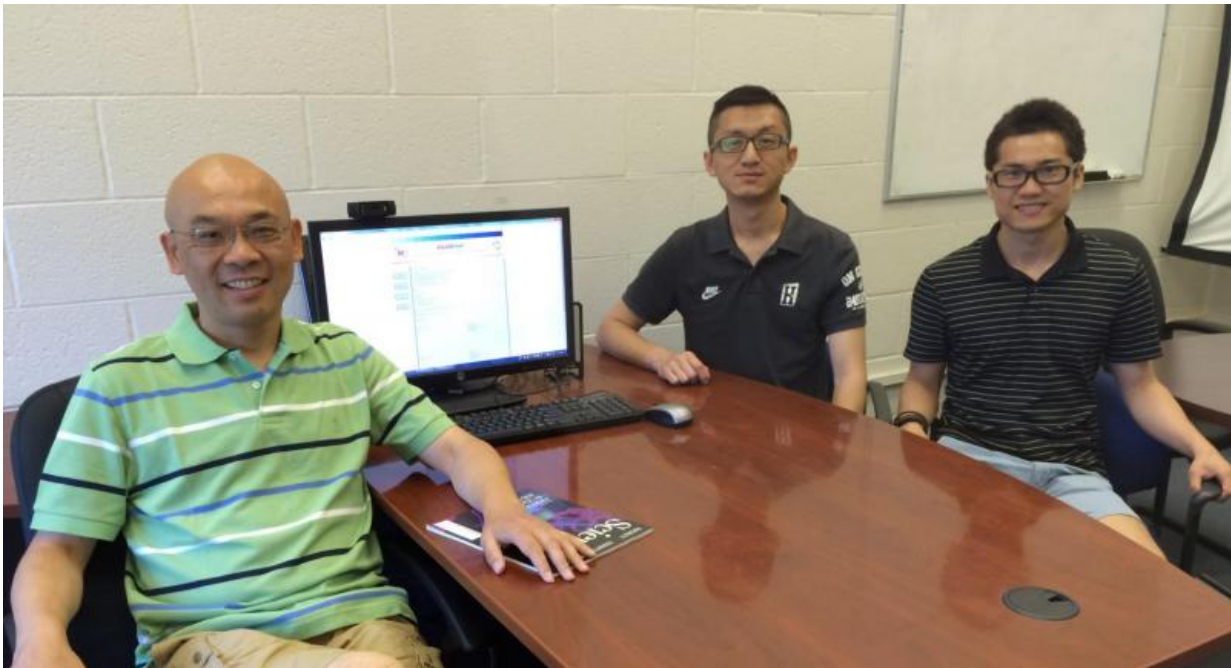# Scientists develop free, online genetic research tool

July 7 2015

Jianlin Cheng, Jilong Li and Jie Hou recently developed RNAMiner, a website making it easier for scientists to analyze genomic data. Credit: Ryan Owens, MU College of Engineering

Technology rapidly is advancing the study of genetics and the search for causes of major diseases. Analysis of genomic sequences that once took days or months now can be performed in a matter of hours. Yet, for most genetic scientists, the lack of access to computer servers and

programs capable of quickly handling vast amounts of data can hinder genetic advancements. Now, a group of scientists at the University of Missouri has introduced a game changer in the world of biological research. The online, free service, RNAMiner, has been developed to handle large datasets which could lead to faster results in the study of plant and animal genomics.

"This work actually started mainly because of the demand of MU scientists," said Jianlin Cheng, an associate professor of computer science in the MU College of Engineering. "RNA sequencing is the means by which researchers use modern sequencing techniques to study RNA, or ribonucleic acid. The process has increased the speed that researchers can note the differences in gene expression among genomes—but it comes at a cost. Often, scientists must sift through incredibly large amounts of data to get to usable results. RNAMiner has cut that time drastically."

Cheng and doctoral students Jilong Li and Jie Hou partnered with members of the MU Center for Botanical Interaction Studies, the Division of Biological Sciences, the Department of Chemistry, the Department of Biochemistry, the MU Informatics Institute and the Bond Life Sciences Center to analyze vast genomic data sets and to formulate the design of RNAMiner.

The website was created to be user-friendly and allows users to upload data, analyze it through as many as five steps against the complete genomes of five species: human, mouse, Drosophila melanogaster (a type of fly), TAIR10 arabidopsis (a small flowering plant) and Clostridium perfringens (a type of bacterium). Genomic data for any species is welcome for upload to grow the database.

On average, two gigabytes of data takes approximately 10 hours for the servers to process and analyze. Most researchers get results within a

couple of hours, Cheng said.

The online, free service, RNAMiner, has been developed to handle large data sets which could lead to faster results in the study of plant and animal genomics. Credit: Jianlin Cheng

"To use our pipeline, you don't have to know about computing tools," Cheng said. "You just need to upload files and select several parameters, and it will automatically give those results. Using this raw data, we can compress that basically hundreds of thousands of times, even one million times, and make the connections needed for our collaborators to identify the genes that cause diseases or certain traits of plants and do some experiments to verify their findings."

More information: The website is free to use and can be found here: calla.rnet.missouri.edu/rnaminer/index.html

Provided by University of Missouri-Columbia