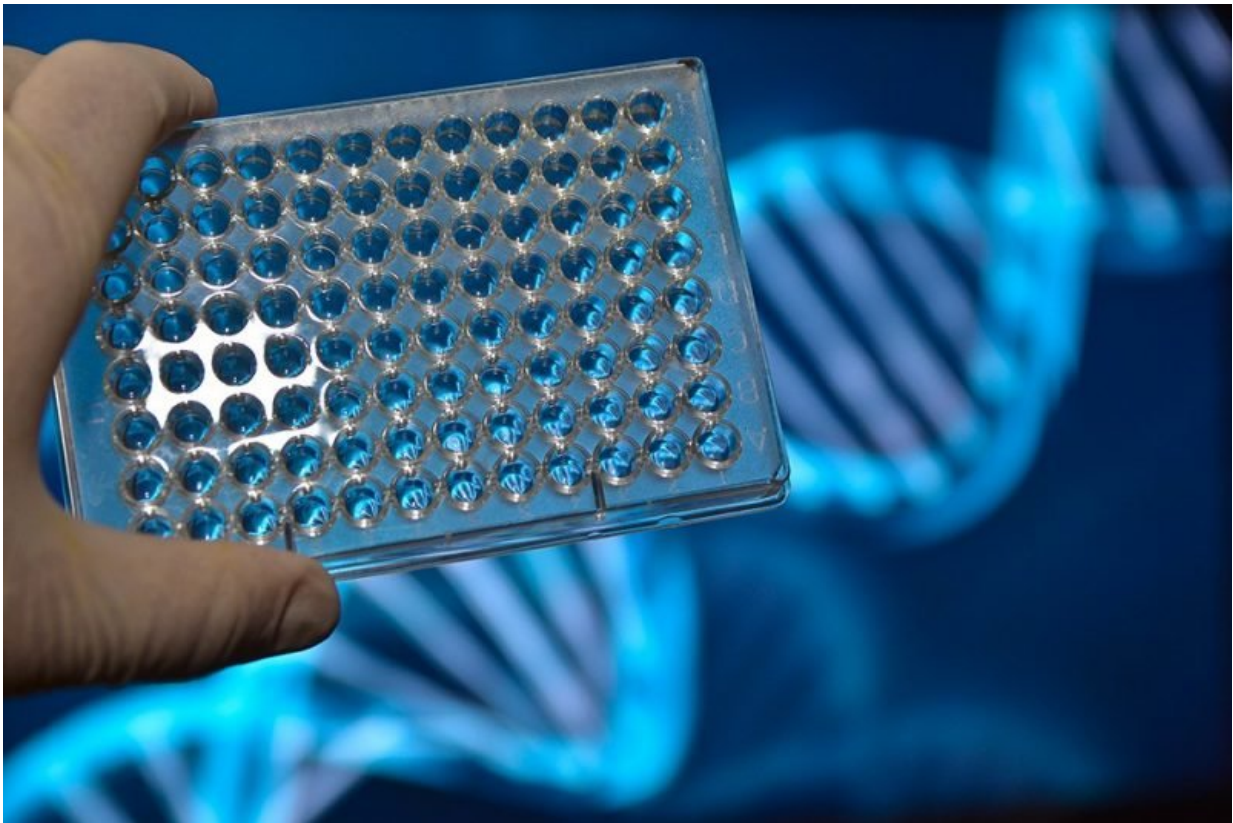


Research team creates powerful system to identify biological threats

November 6 2017, by Tiffany Trent



The Functional Genomic and Computational Assessment of Threats (FunGCAT) program challenges research teams to develop new approaches for screening genetic information for evidence of a potential biological threat. Credit: Virginia Tech

A team led by investigators from the Biocomplexity Institute of Virginia

Tech has been selected through a competitive process to participate in a multimillion dollar program sponsored by the Intelligence Advanced Research Projects Activity (IARPA).

The [Functional Genomic and Computational Assessment of Threats](#) (Fun GCAT) program challenges research teams to develop new approaches and tools for screening nucleic acid sequences and for annotating and characterizing genes of concern, with the goal of preventing accidental or intentional creation of a biological threat.

"Until recently, we could hardly detect [biologically active molecules](#) in our environment," said Stephen Eubank, deputy director of the Network Dynamics and Simulation Science Laboratory and principal investigator on the project. "Now, we know how to synthesize them from scratch. The Fun CGAT project will allow us to ensure we don't create dangerous biological materials."

In only 18 months, the institute team will create a system that predicts the potential function, taxonomy, and danger of unknown nucleotide sequences. The system will function much like an [iNaturalist app](#) for genomic data, with the added feature of predicting how the combinations of DNA and RNA might be used. Using advanced computational and machine learning methods, like neural networks, researchers will collect, standardize, and categorize many types of data in a single flexible, scalable system.

The Biocomplexity Institute's expertise in the Fun GCAT domain stems from a long history of building such systems. In particular, the PAThosystems Resource Integration Center, known as PATRIC, has standardized annotations for almost 110,000 genomes, largely of infectious diseases, over the past decade.

PATRIC is a collaboration between the University of Chicago, the

Biocomplexity Institute, and the Fellowship for Interpretation of Genomes, funded by the National Institute of Allergy and Infectious Disease under the National Institutes of Health. As one of four Bioinformatics Resource Centers, PATRIC aids biomedical researchers through integration of vital pathogen information with rich data and analysis tools.

Rick Stevens, principal investigator of the PATRIC project, said, "The FunGCAT project will foster advancements in machine learning and other computational techniques to enable functional prediction capabilities beyond current traditional bioinformatic techniques. The development work done in PATRIC and other Bioinformatics Resource Centers serves as a prerequisite for attempting these types of predictions." Stevens is also associate director for computing, environment, and life sciences at Argonne National Laboratory.

Systems developed by each of four teams chosen by IARPA will be assessed on speed, accuracy, precision, and deployability. The best systems will be invited to continue for another one to two years of additional funding.

"We have to develop a system that structures the world's knowledge about biological sequences and molecular function so that it can be targeted to answer this challenge," said Andrew Warren, senior research associate at the Biocomplexity Institute. "It's valuable to receive funding that targets fundamental research in this way."

The Biocomplexity Institute's team includes computational biologists, computer scientists, statisticians, and epidemiologists centered mostly at the institute but encompassing other departments at Virginia Tech such as the Discovery Analytics Center and the departments of Statistics and Computer Science, as well as institutions across the country like Icahn School of Medicine at Mount Sinai, Purdue University, Iowa State

University, and San Diego State University.

Gaurav Pandey of the Icahn School of Medicine at Mount Sinai, an expert in ensemble methods and a member of the team, said, "In addition to leveraging computational models to make predictions from a variety of data types, it is also important for the system's performance to assimilate the knowledge embedded in these models. Innovative ensemble methods that embrace the diversity or heterogeneity of these models will help us achieve that goal."

Madhav Marathe, director of the Network Dynamics and Simulation Science Laboratory, said, "IARPA is well known for challenging researchers to take on hard technical problems with important societal impacts. We are excited to develop approaches rooted in high-performance data science, artificial intelligence, and machine learning to address this important health science problem." Under the auspices of the Fun GCAT program, Biocomplexity Institute researchers and their collaborators are doing just that.

Provided by Virginia Tech

Citation: Research team creates powerful system to identify biological threats (2017, November 6) retrieved 21 June 2024 from <https://phys.org/news/2017-11-team-powerful-biological-threats.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.