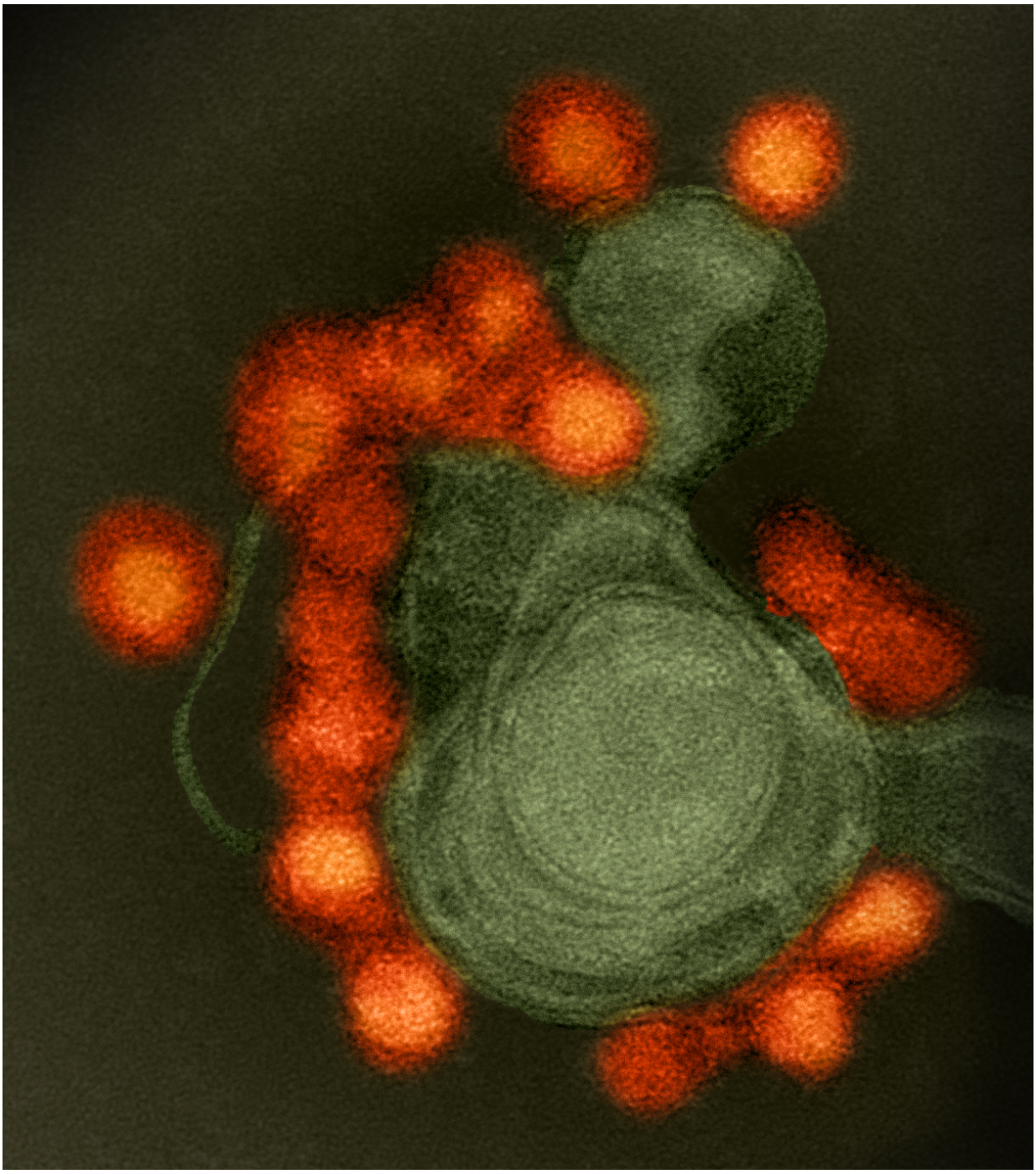
Machine learning tool can predict viral reservoirs in the animal kingdom

November 2 2018



Transmission electron microscope image of negative-stained, Fortaleza-strain Zika virus (red), isolated from a microcephaly case in Brazil. The virus is associated with cellular membranes in the center. Credit: NIAID

Many deadly and newly emerging viruses like Ebola and Zika circulate in wild animal and insect communities long before spreading to humans and causing severe disease. However, finding these natural virus hosts – which could help prevent the spread to humans – currently poses an enormous challenge for scientists.

Now, a new machine learning algorithm has been designed to use viral genome sequences to predict the likely natural host for a broad spectrum of RNA viruses, the viral group that most often jumps from animals to humans.

The new research, led by the University of Glasgow and published today in *Science*, suggests this new tool could help inform preventive measures against deadly diseases. Scientists now hope this new machine learning tool will accelerate research, surveillance and disease control activities to target the right species in the wild, with the ultimate aim of preventing deadly and dangerous viruses reaching humans.

Finding animal and insect hosts of diverse viruses from their genome sequences can take years of intensive field research and laboratory work. The delays caused by this mean that it is difficult to implement preventive measures such as vaccinating the animal sources of disease or preventing dangerous contact between species.

Researchers studied the genomes of over 500 viruses to train machine learning algorithms to match patterns embedded in the viral genomes to their animal origins. These models were able to accurately predict which animal reservoir host each virus came from, whether the virus required the bite of a blood-feeding vector and, if so, whether the vector is a tick, mosquito, midge, or sandfly.

Next, researchers applied the models to viruses for which the hosts and vectors are not yet known, such as Crimean Congo Hemorrhagic Fever,

Zika and MERS. Model predicted hosts often confirmed the current best guesses in each field.

Surprisingly though, two of the four species of Ebola which were presumed to have a bat reservoir, actually had equal or stronger support as primate viruses which could point to a non-human primate, rather than bat, source of some Ebola outbreaks.

Dr. Daniel Streicker, the senior author of the study from the MRC-University of Glasgow Centre for Virus Research, said: "Genome sequences are just about the first piece of information available when viruses emerge, but until now they have mostly been used to identify viruses and study their spread.

"Being able to use those genomes to predict the natural ecology of viruses means we can rapidly narrow the search for their animal reservoirs and vectors, which ultimately means earlier interventions that might prevent viruses from emerging all together or stop their early spread."

Dr. Pete Gardner from Wellcome's Infection & Immunobiology team said: "Healthy animals can carry viruses which can infect people causing disease outbreaks. Finding the animal species is often incredibly challenging, making it difficult to implement preventative measures such as vaccinating animals or preventing animal contact.

"This important study highlights the predictive power of combining machine learning and genetic data to rapidly and accurately identify where a disease has come from and how it is being transmitted. This new approach has the potential to rapidly accelerate future responses to viral outbreaks."

The researchers are now developing a web application that will allow

scientists from anywhere in the world to submit their virus sequences and get rapid predictions for reservoir hosts, vectors and transmission routes.

More information: Simon A. Babayan et al. Predicting reservoir hosts and arthropod vectors from evolutionary signatures in RNA virus genomes, *Science* (2018). DOI: 10.1126/science.aap9072

Provided by University of Glasgow

Citation: Machine learning tool can predict viral reservoirs in the animal kingdom (2018, November 2) retrieved 19 April 2024 from https://phys.org/news/2018-11-machine-tool-viral-reservoirs-animal.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.