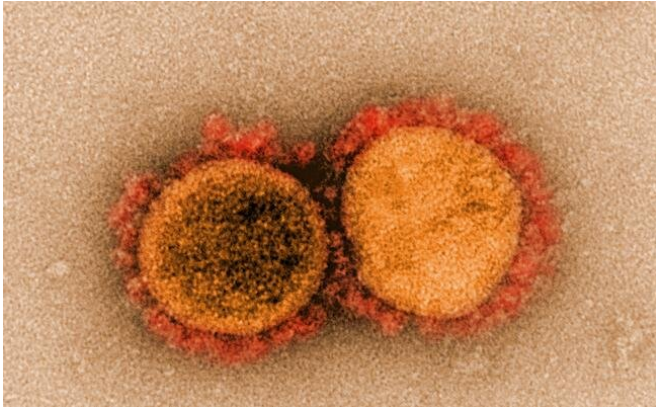


Viral sequencing can reveal how SARS-CoV-2 spreads and evolves

28 January 2021



Transmission electron micrograph of SARS-CoV-2 virus particles, isolated from a patient. Image captured and color-enhanced at the NIAID Integrated Research Facility (IRF) in Fort Detrick, Maryland. Credit: NIAID

The emergence of SARS-CoV-2 virus variants that are adding twists in the battle against COVID-19 highlight the need for better genomic monitoring of the virus, says Katia Koelle, associate professor of biology at Emory University.

"Improved genomic surveillance of SARS-CoV-2 across states would really help us to better understand how the [virus](#) causing the pandemic is evolving and spreading in the United States," Koelle says. "More federal funding is needed, along with centralized standards for [sample collection](#) and genetic sequencing. Researchers need access to such metadata to better track how the virus is spreading geographically, and to identify any new variants that may make it harder to control, so that health officials can respond more quickly and effectively."

Koelle studies the interplay between [viral evolution](#) and the epidemiological spread of viral infectious diseases. She is senior author of a "Viewpoint" article just published in the journal *Science* on the

importance of SARS-CoV-2 sequencing to control the COVID-19 pandemic.

Michael Martin, a Ph.D. student in Emory's Population, Biology and Ecology Program and a member of Koelle's lab, is first author of the Science article. David VanInsberghe, a post-doctoral fellow in Koelle's lab, is co-author.

"Research into SARS-CoV-2 has been going at lightning speed," Martin says. "This acceleration has provided us with one of the largest datasets ever so quickly assembled for a disease. We've learned a lot so far about how this virus spreads and adapts, but we still have many blind spots that need to be addressed."

The article summarizes key insights about SARS-CoV-2 that have already been gained by sequencing of its genome from individual patient samples. It also cites challenges that remain, including the collection and integration of metadata into genetic analyses and the need for the development of more efficient and scalable computational methods to apply to hundreds of thousands of genomes.

A genome is an organism's genetic material. Human genomes are made up of double-stranded DNA, coded in four different nucleotide base letters. A single human genome consists of more than 3 billion base pairs. In contrast, the genome of coronaviruses, including SARS-CoV-2, are made of RNA, which can have a simpler structure than DNA. The SARS-CoV-2 genome, for instance, consists of a single RNA strand that is only 30,000 letters long. Sequencing is a technique that provides a read-out of these letters.

If the SARS-CoV-2 virus is found in a sample swabbed from someone's nose or mouth, it confirms the likelihood that the person is carrying the virus, whether they have symptoms of COVID-19 or not. The virus in the sample can also

be sequenced.

"Sequencing the virus is like fingerprinting it," Koelle explains. "And based on how close the fingerprints match between samples—that is, how close they are genetically—you can at times learn who is infecting whom. Analyzing sequences from samples taken from infected individuals in a given region over time can provide even more information."

Analyses of SARS-CoV-2 sequencing data have enabled researchers to estimate the timing of SARS-CoV-2 spillover into humans; identify some of the transmission routes in its global spread; determine infection rates and how they change within a region; and identify the emergence of some new variants of concern.

Viral genomes can mutate during replication, changing letters as they spread to new people. Most of these random mutations will likely not affect the transmissibility or virulence of a virus—but a few may make it even more difficult to fight. Early evidence, for instance, suggests that a SARS-CoV-2 variant that recently emerged in the UK may be more easily transmitted and potentially more severe. A South African variant shows signs that it may reduce the efficacy of existing vaccines, while a variant first detected in Brazil also contains mutations that health officials worry may make the virus spread more quickly.

"It can be difficult to identify which variants actually change how the virus replicates, spreads and causes disease because of confounding factors," Martin explains. "If a variant spreads more quickly, for instance, you have to tease apart whether that was due to it becoming more transmissible or if someone who was infected with it attended a large gathering."

The better data researchers have, the faster they can solve such puzzles, he adds.

Technological advances during recent years have made it more efficient and less costly to generate sequencing data. Barely a year after it emerged, more than 400,000 sequences of SARS-CoV-2 are now available in public databases, such as the

GISAID platform which was launched in 2008 to share information among National Influenza Centers for the WHO Global Influenza Surveillance and Response System.

"A large chunk of the public sequencing data for SARS-CoV-2 has come out of the UK," Koelle notes. "That's because the British government has an initiative to do high-density sampling of the SARS-CoV-2 genome."

The rich data set from the UK helped identify the emergence of the variant in Britain that is spreading rapidly. "There might be other variants of concern emerging in other places around the world besides the ones already identified, but we just don't know because we don't have as good of surveillance in those locations," Koelle says.

"While the United States has been slow in efforts to sequence SARS-CoV-2 from samples across the nation, there are several excellent viral sequencing efforts and phylogenetic analyses, primarily driven by academic researchers, that have helped to understand SARS-CoV-2 transmission more locally," Koelle says. "We have the expertise in the U.S., but the effort is more piecemeal."

"We need a coordinated, nationally standardized program to do widespread sequencing of SARS-CoV-2 in the United States," Martin says. "Much of the data collected now just has a state identifier but we need greater resolution while also protecting patient privacy. More county-level identifiers, for instance, would be one way to greatly improve the quality and the depth of the data."

Once the COVID-19 pandemic ebbs, it's important to continue to build the national infrastructure and systems for infectious disease surveillance—including viral sequencing—and to keep it in place, both researchers stress.

"There will be more infectious disease pandemics, and we need to be better prepared," Martin says.

More information: M.A. Martin et al., "Insights from SARS-CoV-2 sequences," *Science* (2020). science.sciencemag.org/cgi/doi/10.1126/science.abf3995

Provided by Emory University

APA citation: Viral sequencing can reveal how SARS-CoV-2 spreads and evolves (2021, January 28)
retrieved 27 April 2021 from

<https://medicalxpress.com/news/2021-01-viral-sequencing-reveal-sars-cov-evolves.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.