

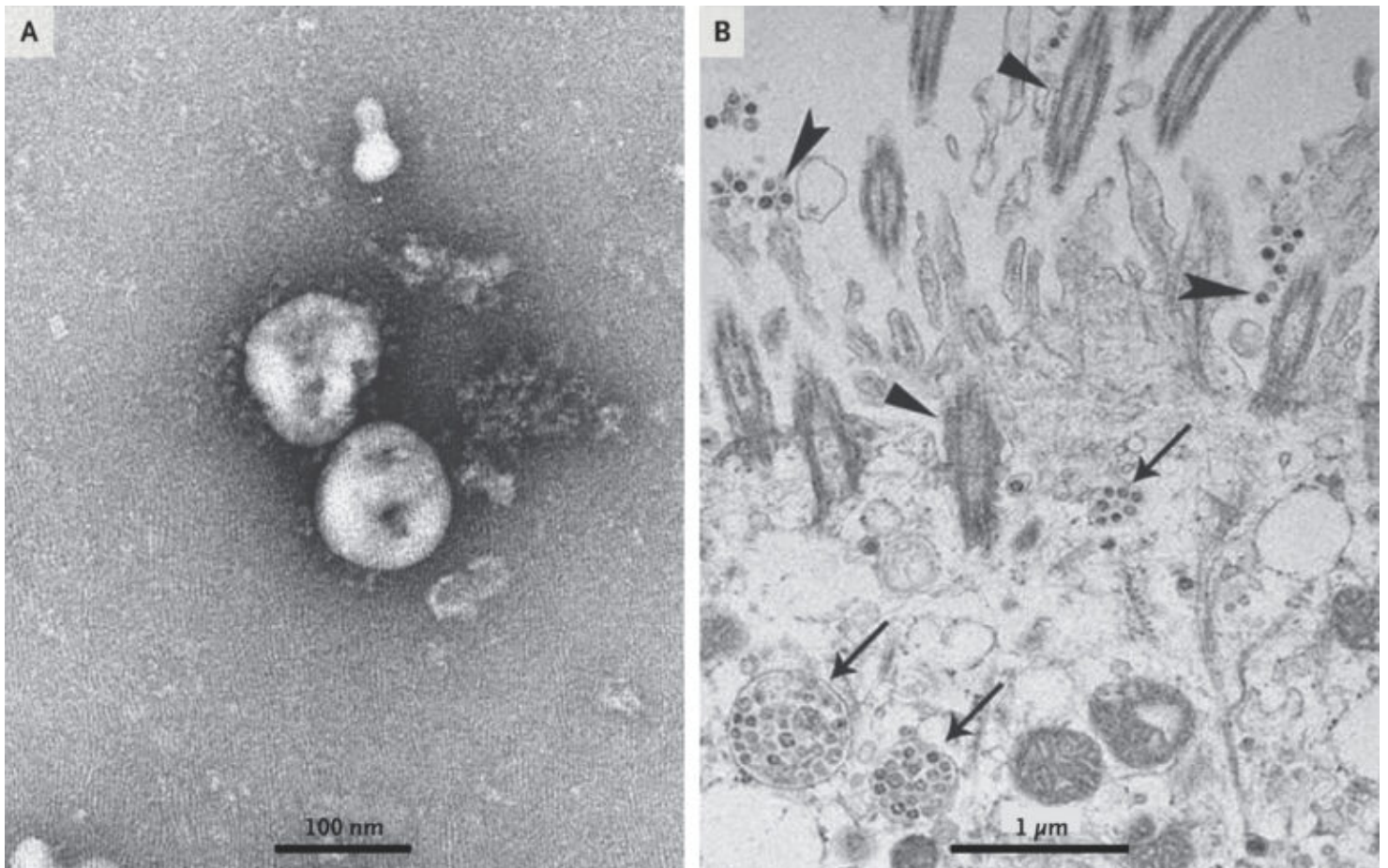
STAT

1

DNA sleuths read the coronavirus genome, tracing its origins and looking for dangerous mutations

By [Sharon Begley](#)² [@sxbegle](#)³

January 24, 2020



Electron micrographs of isolated 2019-nCoV particles (left), and in cells from human airways (right), marked with arrows. *The New England Journal of Medicine*

As infectious disease specialists and epidemiologists race to contain the outbreak of the novel coronavirus centered on Wuhan, China, they're getting backup that's been possible only since the explosion in genetic technologies: a deep-dive into the genome of the virus known as 2019-nCoV.

Analyses of the viral genome are already providing clues to the origins of the outbreak and even possible ways to treat the infection, a need that is becoming more urgent by the day: Early on Saturday in China, health officials reported 15 new fatalities in a single day, bringing the death toll to 41. There are now nearly 1,100 confirmed cases there.

Reading the genome (which is made of RNA, not DNA) also allows researchers to monitor how 2019-nCoV is changing and provides a roadmap for developing a diagnostic test and a vaccine.

“The genetics can tell us the true timing of the first cases” and whether they occurred earlier than officials realized, said molecular biologist Kristian Andersen of Scripps Research, an expert on viral genomes. “It can also tell us how the outbreak started — from a single event of a virus jumping from an infected animal to a person or from a lot of animals being infected. And the genetics can tell us what’s sustaining the outbreak — new introductions from animals or human-to-human transmission.”

Scientists in China sequenced the virus’s genome and made it [available](#)⁴ on Jan. 10, just a month after the Dec. 8 report of the first case of pneumonia from an unknown virus in Wuhan. In contrast, after the SARS outbreak began in late 2002, it took scientists much longer to sequence that coronavirus. It peaked in February 2003 — and the complete genome of 29,727 nucleotides wasn’t [sequenced](#)⁵ until that April.

Support STAT: ⁶STAT is offering coverage of the coronavirus for free. Please consider a subscription to support our journalism. Start free trial today.⁷

Since the sequencing of the first 2019-nCoV sample, from an early patient, scientists have completed nearly two dozen more, said Andrew Rambaut of the University of Edinburgh, an expert on viral evolution. That pace is

“unprecedented and completely unbelievable,” said Andersen, who worked on sequencing the Ebola genome during the 2014 outbreak. “It’s just insane.”

The genome of the Wuhan virus is 29,903 bases long, one of many clues that have led scientists to believe it is very similar to SARS.

By comparing the two dozen genomes, scientists can address the “when did this start” question. The 24 available samples, including from Thailand and Shenzhen as well as Wuhan, show “very limited genetic variation,” Rambaut concluded on an online discussion forum where virologists have been sharing data and analyses. “This is indicative of a relatively recent common ancestor for all these viruses.”

Given what’s known about the pace at which viral genomes mutate, if nCoV had been circulating in humans since significantly before the first case was reported on Dec. 8, the 24 genomes would differ more. Applying ballpark rates of viral evolution, Rambaut estimates that the Adam (or Eve) virus from which all others are descended first appeared no earlier than Oct. 30, 2019, and no later than Nov. 29.

The progenitor virus itself was almost certainly one that circulates harmlessly in bats (as SARS does) but has an “intermediate reservoir” in one or more animals that come into contact with people, Andersen said. Presumably, that reservoir is one of the species of animals at the Wuhan market thought to be ground zero for the outbreak. The ancestor of 2019-nCoV existed in that species for some unknown time, never infecting people, until by chance a single virus acquired a mutation that made it capable of jumping into and infecting humans.

The genome sequences suggest that was a one-time-only jump. “The genomes [from the 24 samples] are very uniform,” Andersen said. “If there had been multiple introductions,” including from many different animals, “there would be more genomic diversity. This was a single introduction.”

That means that what's sustaining the spread is human-to-human transmission (suggesting that closing Wuhan's animal market is very much an after-the-horse-has-fled-the-barn reaction).

[Related:](#) ¹⁰

[**New coronavirus can cause infections with no symptoms and sicken otherwise healthy people, studies show**](#) ¹⁰

Unfortunately, genetic analysis can't identify what animal species the coronavirus jumped from into humans. But an analysis by a team from the Wuhan Institute of Virology, [posted](#) ¹¹ to the preprint server bioRxiv, determined that the genome of this coronavirus (the seventh known to infect humans) is 96% identical to that of a bat coronavirus, suggesting that species is the original source. (Writing in the New England Journal of Medicine on Friday, another team of scientists in China [reported](#) ¹² that the new coronavirus is 86.9% identical to the bat SARS-like coronavirus.)

Virologists differ on whether it's possible to read out viral properties from just the genome sequence, such as whether the microbe is spread by coughing, sneezing, touching, or merely breathing. But the analysis by the Wuhan Institute team found that it enters human cells using the same doorway that SARS did. Called angiotensin converting enzyme 2 (ACE2), the door is a receptor to which a "spike protein" on the virus's surface first attaches and then enables the virus to fuse with the host cell.

If ACE2 is "druggable," blocking it could conceivably treat 2019-nCoV. "It should be expected and worth to test if ACE2 targeting ... drugs can be used for nCoV-2019 patients," the scientists wrote.

The genome sequences have more to give. They "will be crucially important for development of diagnostics [and] vaccines," said biologist Richard Ebright of Rutgers University.

For instance, the genome-editing technology CRISPR is the basis for Cambridge, Mass.-based [startup Sherlock Biosciences](#)'¹³ diagnostics, which promise to slash how long it takes to make a definitive identification. In the U.S, that's now done only by sending samples to the Centers for Disease Control and Prevention, which uses a technology invented in the 1980s, polymerase chain reaction or PCR, to identify the presence of coronavirus.

“Our vision is that our [CRISPR-based] SHERLOCK and INSPECTR platforms are tailor-made for outbreaks like coronavirus,” said Sherlock CEO Rahul Dhanda, who declined to discuss “specific plans related to coronavirus.”

And as scientists keep adding 2019-nCoV genome sequences to their collection, they could get an early glimpse of whether the virus is mutating in a way that could make it more dangerous or more transmissible. “You need continuous sequencing,” Andersen said.

Correction: This story has been corrected to make clear that the coronavirus genome is made of RNA, not DNA.

About the Author



[Sharon Begley](#)²

Senior Writer, Science and Discovery

Sharon covers science and discovery.

sharon.begley@statnews.com¹⁴

[@sxbege](#)³

Tags

Links

1. <https://www.statnews.com/category/in-the-lab/>
2. <https://www.statnews.com/staff/sharon-begley/>
3. <https://twitter.com/sxbegle>
4. <https://www.ncbi.nlm.nih.gov/nuccore/MN908947>
5. <https://www.cdc.gov/sars/lab/sequence.htm>
6. https://www.statnews.com/stat-plus/?utm_source=website&utm_campaign=coronavirus&utm_medium=w
7. https://www.statnews.com/subscribe/?utm_source=website&utm_campaign=coronavirus&utm_medium=w%20
8. <https://www.statnews.com/signup/>
9. <https://www.statnews.com/privacy/>
10. <https://www.statnews.com/2020/01/24/coronavirus-infections-no-symptoms-lancet-studies/>
11. <https://www.biorxiv.org/content/10.1101/2020.01.22.914952v2>
12. https://www.nejm.org/doi/full/10.1056/NEJMoa2001017?query=featured_home
13. <https://sherlock.bio/>
14. <https://www.statnews.com/2020/01/24/dna-sleuths-read-coronavirus-genome-tracing-origins-and-mutations/comment-page-1/mailto:sharon.begley@statnews.com>
15. <https://www.statnews.com/tag/boston/>
16. <https://www.statnews.com/tag/china/>
17. <https://www.statnews.com/tag/coronavirus/>
18. <https://www.statnews.com/tag/crispr/>
19. <https://www.statnews.com/tag/genetics/>
20. <https://www.statnews.com/tag/global-health/>
21. <https://www.statnews.com/tag/infectious-disease/>

© 2020 STAT