



This Sunda pangolin found throughout Southeast Asia is currently considered to be critically endangered. Piekfrosch / German Wikipedia, CC BY-SA

Study shows pangolins may have passed new coronavirus from bats to humans

Published: April 10, 2020 1.17pm BST

Yang Zhang

Professor of Computational Medicine & Bioinformatics, University of Michigan

Chengxin Zhang

PhD Candidate in Bioinformatics, University of Michigan

Wei Zheng

Postdoctoral Fellow of Computational Medicine and Bioinformatics, University of Michigan

Pangolins, not snakes, may be the missing link for transmission of the new coronavirus from bats to humans.

Since its initial outbreak at Huanan Seafood Wholesale Market in Wuhan, China, in late 2019, COVID-19 has since infected more than a million people across the globe. To understand and control the transmission of COVID-19, scientists are racing to study the coronavirus causing the disease: SARS-CoV-2, previously named 2019-nCoV. SARS-CoV-2 is zoonotic, which means that the virus originated in animals and jumped to humans. A critical challenge is to determine which animal transmitted the virus to humans.

We are a team of bioinformaticians and we feel it is our responsibility to the global community to investigate the origin of this virus.

Based on the research in our lab, we believe that pangolins, as opposed to snakes, may have served as the hosts that transmitted the coronavirus to people and caused the ongoing COVID-19 pandemic. The pangolin, also known as a scaly anteater, is the only known mammal with scales and is found in Asia and Africa.

Mystery of zoonotic transmission

Since January 2020, the current consensus among the scientific community is that SARS-CoV-2 originated in horseshoe bats; however, it's unlikely that bats directly gave the virus to humans based on what's known about transmission of earlier zoonotic coronaviruses. Instead, scientists suspected that the bat coronavirus infected another animal, an "intermediate host," which subsequently transmitted the virus to humans.

For example, SARS-CoV, which is the coronavirus that caused the severe acute respiratory syndrome (SARS) pandemic in 2003, is a close relative of SARS-CoV-2. It was also found to have been transmitted from bats to an intermediate host – the masked palm civet – which subsequently infected humans. Similarly, MERS-CoV, the coronavirus that caused Middle East respiratory syndrome (MERS) in 2012, jumped from bats to another intermediate host, the dromedary camel, before infecting humans.

The identity of the intermediate host of SARS-CoV-2 is therefore a mystery that many researchers hope to solve, as knowing the intermediate host is very helpful for prevention of further spread of epidemic. An early study claimed that snakes such as the Chinese krait and the Chinese cobra were likely to be the intermediate hosts for SARS-CoV-2. Yet, this conclusion quickly drew skepticism, partly because there exists no previous evidence that coronaviruses can jump from a cold-blooded animal, such as snakes, to human beings.

Snakes make unlikely host

The early claim that snakes transmitted SARS-CoV-2 was based on an analysis of the virus's genetic sequence. For both viruses and animal cells to function, genetic sequences (RNA or DNA) must be translated into proteins, which then carry out many tasks of the virus and the cell. These proteins exist as linked chains of single amino acids; each amino acid in a protein is encoded by a group of three nucleotides, also known as a codon, in the genetic sequence.

Since there are 64 possible different codons but only 20 amino acids, several codons can correspond to the same amino acid; different organisms will have different preference for which codon is used for a given amino acid. The early study hypothesized that for the coronavirus to effectively grow inside an animal cell, the codon usage preferences of the coronavirus should match that of the host cell.

The researchers compared the codon usage in the SARS-CoV-2 virus against that of the cells in eight animals at the Wuhan Huanan Seafood Wholesale Market. That study found that the snakes share the most similar codon usage pattern to SARS-CoV-2, thereby declaring that snakes were the most likely intermediate hosts.

However, their central hypothesis that coronaviruses and their animal hosts share similar codon usage was never verified. Our team at the University of Michigan scrutinized this hypothesis, and performed a more systematic analysis that we published in a recent follow-up study. We compared the codon usages of three coronaviruses (SARS-CoV-2, SARS-CoV and MERS-CoV) to those of more than 10,000 different kinds of animals.

To our surprise, we found that the codon usage of a coronavirus is not determined by its hosts. For example, the codon usage of SARS-CoV and MERS-CoV is much closer to frogs and snakes than to their real animal hosts (civets and camels, respectively). This shows that it is not possible to use only codon usage in animals' cells to infer the hosts of coronaviruses, suggesting that the early claim of snake-borne transmission of SARS-CoV-2 is likely to be incorrect.

Discovering the pangolin as a likely missing link

Our follow-up study also found that the genetic sequence of a coronavirus, discovered in lung samples of Malayan pangolins, was highly similar to SARS-CoV-2. The two viruses shared 91% of their genetic sequence.

There is a particularly strong similarity between the spike proteins of these two viruses. The spike protein, which is on the surface of a coronavirus, is used by the virus to get into an animal cell. The bat coronavirus, which was the ancestor of SARS-CoV-2, has 19 amino acids on the spike protein that are different from SARS-CoV-2; the pangolin coronavirus only has five amino acids that are different from SARS-CoV-2. Meanwhile, several other research groups have found further experimental evidence of pangolins being infected by coronaviruses highly similar to SARS-CoV-2.