A close relative of SARS-CoV-2 found in bats offers more evidence it evolved naturally

11 May 2020

There is ongoing debate among policymakers and the general public about where SARS-CoV-2, the virus that causes COVID-19, came from. While researchers consider bats the most likely natural hosts for SARS-CoV-2, the origins of the virus are still unclear. On May 10 in the journal *Current Biology*, researchers describe a recently identified bat coronavirus that is SARS-CoV-2's closest relative in some regions of the genome and which contains insertions of amino acids at the junction of the S1 and S2 subunits of the virus's spike protein in a manner similar to SARS-CoV-2. While it's not a direct evolutionary precursor of SARS-CoV-2, this new virus, RmYN02, suggests that these types of seemingly unusual insertion events can occur naturally in coronavirus evolution, the researchers say.

"Since the discovery of SARS-CoV-2 there have been a number of unfounded suggestions that the virus has a laboratory origin," says senior author Weifeng Shi, director and professor at the Institute of Pathogen Biology at Shandong First Medical University in China. "In particular, it has been proposed the S1/S2 insertion is highly unusual and perhaps indicative of laboratory manipulation. Our paper shows very clearly that these events occur naturally in wildlife. This provides strong evidence against SARS-CoV-2 being a laboratory escape."

The researchers identified RmYN02 from an analysis of 227 bat samples collected in Yunnan province, China, between May and October of 2019. "Since the discovery that bats were the reservoir of SARS coronavirus in 2005, there has been great interest in bats as reservoir species for infectious diseases, particularly as they carry a very high diversity of RNA viruses, including coronaviruses," Shi says. RNA from the samples was sent for metagenomic next-generation sequencing in early January 2020, soon after the discovery of SARS-CoV-2.

Across the whole genome, the closest relative to SARS-CoV-2 is another virus, called RaTG13, which was previously identified from bats in Yunnan province. But RmYN02, the virus newly discovered here, is even more closely related to SARS-CoV-2 in some parts of the genome, including in the longest encoding section of the genome called 1ab, where they share 97.2% of their RNA. The researchers note that RmYN02 does not closely resemble SARS-CoV-2 in the region of the genome that encodes the key receptor binding domain that binds to the human ACE2 receptor that SARS-CoV-2 uses to infect host cells. This means it's not likely to infect human cells.

The key similarity between SARS-CoV-2 and RmYN02, is the finding that RmYN02 also contains amino acid insertions at the point where the two subunits of its spike protein meet. SARS-CoV-2 is characterized by a four-amino-acid insertion at the
junction of S1 and S2; this insertion is unique to the virus and has been present in all SARS-CoV-2 sequenced so far. The insertions in RmYN02 are not the same as those in SARS-CoV-2, which indicates that they occurred through independent insertion events. But a similar insertion event happening in a virus identified in bats strongly suggests that these kinds of insertions are of natural origin. "Our findings suggest that these insertion events that initially appeared to be very unusual can, in fact, occur naturally in animal betacoronaviruses," Shi says.

"Our work sheds more light on the evolutionary ancestry of SARS-CoV-2," he adds. "Neither RaTG13 nor RmYN02 is the direct ancestor of SARS-CoV-2, because there is still an evolutionary gap between these viruses. But our study strongly suggests that sampling of more wildlife species will reveal viruses that are even more closely related to SARS-CoV-2 and perhaps even its direct ancestors, which will tell us a great deal about how this virus emerged in humans."


Provided by Cell Press