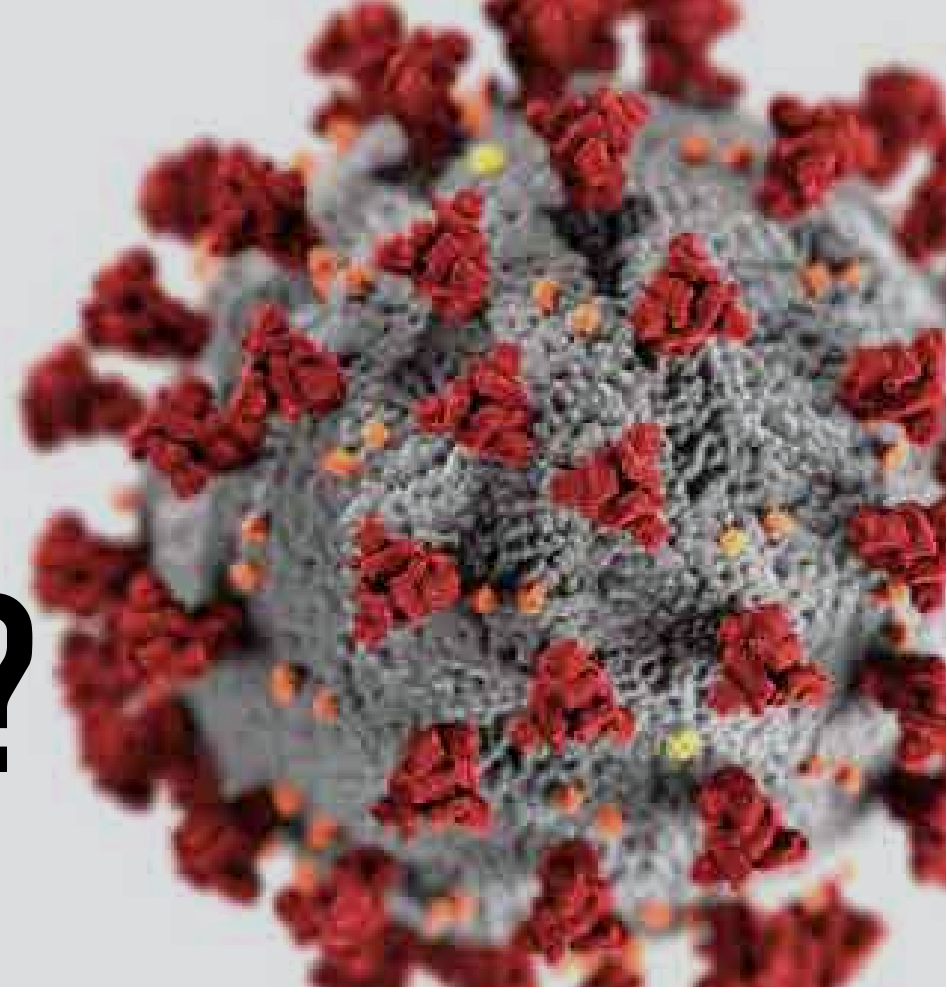
WHAT DO WE KNOW ABOUT SARS-CoV-2?



SHAHID JAMEEL

How do we know that SARS-CoV-2 is a 'novel' coronavirus? How was this virus discovered? What do we know about its structure? How long can it remain infectious outside a host cell? Is there any evidence to suggest that SARS-CoV-2 may have originated in a lab?

For some time now, scientists have predicted the emergence of "Disease X", a hypothetical unknown disease agent that would be capable of causing a future epidemic. In December 2019, the world woke up to the news of a new virus. Initially referred to as the 2019 novel coronavirus, it was later renamed as SARS-CoV-2.

A novel coronavirus

The 'SARS' in its name reflects the symptoms this virus produces in some infected people — Severe Acute Respiratory Syndrome, and 'CoV-2' refers to the fact that it is the second coronavirus found to produce these symptoms (the first one emerged in 2002-03 and was called SARS-CoV). Coronaviruses are a group of viruses with surface spikes, visible under a very high-powered electron microscope, that give them a solar corona or crown-like (hence the name) appearance (see **Fig. 1**). The disease this virus causes is called Coronavirus Infectious Disease 2019, abbreviated as COVID-19.

Coronaviruses cause acute and persistent infections in humans, other mammals, and birds. Members of this family were isolated from animals as early as in the 1930s. But it was only in the 1960s that some viruses causing respiratory disease in humans were discovered to be part of

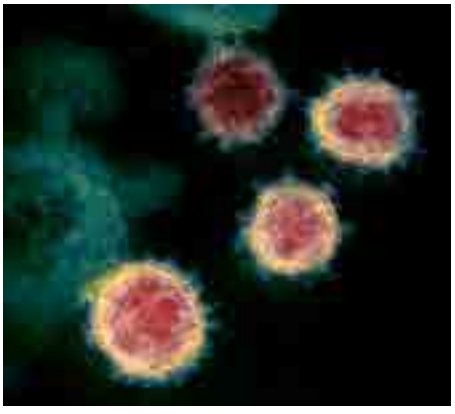


Fig. 1. Electron micrograph of SARS-CoV-2 particles with visible coronae. Virus particles are shown emerging from the surface of cells cultured in the lab.

Credits: NIAID Rocky Mountain Laboratories (RML), U.S. NIH. URL: https://www.flickr.com/photos/niaid/49534865371/in/album-72157712914621487/. License: CC-BY.

this family. Four human coronaviruses were identified as being endemic in the human population. These cause about 20% of common colds annually.

Novel viruses are emerging all the time, but come into prominence only when they cause disease in humans or other animal species. In the past two decades, three new human coronaviruses have emerged, all from bats, to cause outbreaks in different parts of the world. These include the SARS-CoV in 2002-03 in China, the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) in Saudi Arabia in 2012, and the most recent SARS-CoV-2 in China in 2019. All these viruses have similar physical attributes and genome organization, but differ in their genetic sequences. For example, SARS-CoV-2 has ~80% sequence identity with SARS-CoV, but only ~55% to MERS-CoV, and ~50% to the common cold coronaviruses.

What do we know about SARS-CoV-2?

In late 2019, doctors in China's Hubei Province, and its main city Wuhan, started seeing a cluster of patients with pneumonia. This is an infection of the lung in which the air sacs are inflamed and may fill with fluid or pus, causing cough, fever, chills, and breathing difficulties (see **Fig. 2**). While pneumonia is seen in infections by many microorganisms (bacteria, viruses, and fungi), genetic sequencing of fluid from the lungs of patients from Wuhan always showed the presence of virus sequences similar, but not identical, to SARS-CoV. The virus was isolated from cell cultures infected with fluids from the lungs, throat, and nasal cavity of patients. Thus, powerful detection tools, cell culture, and genetic sequencing technologies ensured that the virus was identified within just a few weeks.

Like other coronaviruses, SARS-CoV-2 is a particle of about 100 nanometers (1 nanometer = 10^{-9} meter) in size. This enveloped virus has three different proteins — the Spike (S), Envelope (E), and Membrane (M) proteins — embedded in its lipid (fat) envelope. It is the S proteins that give the virus its corona-like appearance. Within the envelope, a single strand of RNA of about 30,000 bases or nucleotides is tightly wrapped around multiple copies of the Nucleocapsid (N) protein (see **Fig. 3**). During infection, the virus RNA enters a host cell, replicates, and directs the synthesis of about 24-27 proteins that allow the virus to multiply. Some of these proteins, like the viral replicase enzyme and the S, E, M and N proteins, help in direct ways. Several nonstructural proteins, such as the Orf3a protein, help indirectly by modifying cellular processes in the host to better replicate the virus.

While a genome of 30,000 bases in coronaviruses is unusually large for an RNA virus (most are about a third of this), its size is not important for transmission or severity of disease. Large population sizes and a high frequency of errors introduced during genome replication mean that viruses tend to show high mutation rates. In general, RNA viruses mutate even faster than DNA viruses because their replicases (enzymes which catalyze the synthesis of a complementary RNA molecule

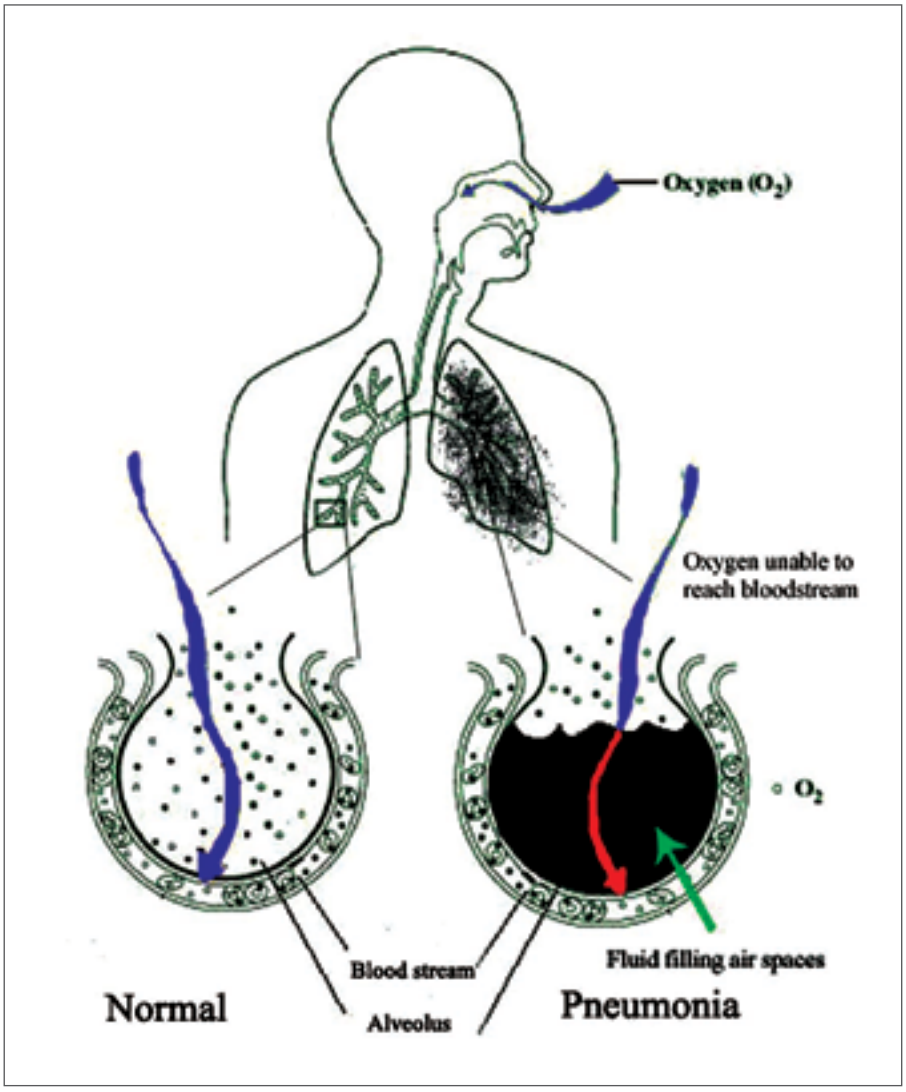


Fig. 2. The discovery of COVID-19 was sparked by a cluster of patients with pneumonia in Wuhan, China. Lung infection can lead to fluid accumulation in air sacs that results in breathing difficulties.

Credits: National Institutes of Health, United States Department of Health and Human Services, Wikimedia Commons. URL: https://commons.wikimedia.org/wiki/File:New_Pneumonia_cartoon.jpg. License: CC-BY.

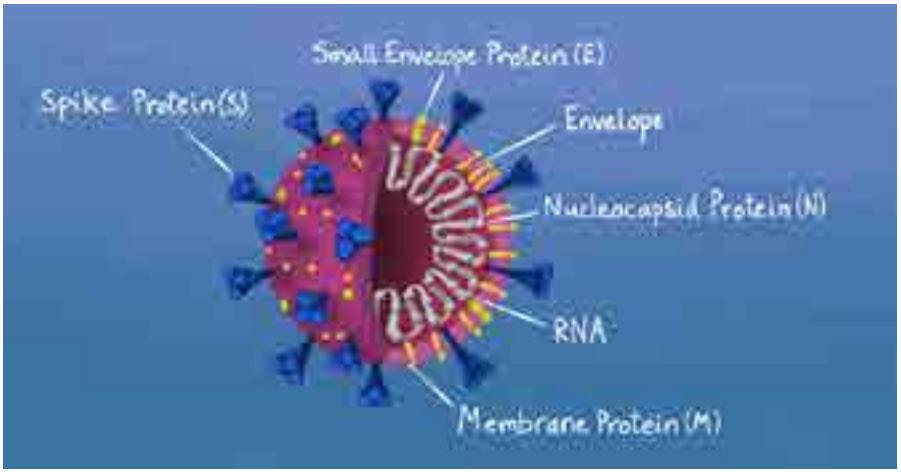


Fig. 3. Artist's model of a SARS-CoV-2 particle. An estimated 100 trimers (or 300 monomers) of S, about 2000 copies of M, and about 20 copies of E are deeply embedded in the lipid envelope of each virus particle. Within the envelope, an RNA strand of 30,000 bases is tightly wrapped around multiple copies of the Nucleocapsid (N) protein.

Credits: Created by Maya Peters Kostman for the Innovative Genomics Institute. URL: https://innovativegenomics.org/wp-content/uploads/2020/04/Single-virion-with-all-parts-labeled.png. License: CC-BY-NC-SA.

using an RNA template) lack an error-correction function. Higher mutation rates lead to more diverse virus populations, some of which randomly acquire the ability to survive or transmit better. However, due to an enzyme that repairs replication errors, coronaviruses tend to have lower mutation rates that are, for example, about 1000 times lower than that for influenza viruses (also RNA viruses).

As of now, only one strain of SARS-CoV-2 is circulating globally. But this strain has multiple groups (or clades) of isolates, which often show variations in geographic distribution. Since its emergence, one mutant of the virus, with an aspartic acid to glycine change at amino acid 614 in a Spike protein, has spread faster than others. Scientists believe that this mutation (called D614G) has increased both the stability of the Spike protein trimers and their ability to bind cell surface receptors for productive infection.

Since a virus is "living" only when inside a host cell, we call it either infectious or noninfectious when outside (see **Box 1**). SARS-CoV-2 can stay infectious on surfaces for variable times at room temperature — up to 4 hours on copper, 24 hours on cardboard, and 72 hours on plastic and stainless steel. In laboratory tests, SARS-CoV-2 can be inactivated within seconds at temperatures of 90°C or above. It is highly stable at 4°C, and stays infectious at 22°C for 7 days, 37°C for 1 day, 56°C for 10 min, and 70°C for 1 min. A 1:100 dilution of household bleach (4% sodium hypochlorite) is helpful in disinfecting surfaces, but should not be used on human skin as it can cause irritation, especially in the eyes.

Origin of SARS-CoV-2

Where did SARS-CoV-2 come from? Evidence indicates that it jumped into humans from bats, either directly or via another animal species (in this case, the pangolin — a scaly mammal whose flesh is eaten in China, and scales are used in Chinese traditional medicine).

What is the evidence? The SARS-CoV-2 genome has 96% sequence identity to coronaviruses isolated from bats in eastern China in 2018, and 91% identity to coronaviruses isolated from pangolins. This is much higher than its genetic identity to SARS-CoV (~80%), MERS-CoV (~55%) and other human coronaviruses (~50%), suggesting that SARS-CoV-2 is not derived from the earlier human viruses. Two key features in the Spike proteins of SARS-CoV-2 allow it to bind human cells and transmit efficiently. One of these is found in Pangolin-CoV, but both are missing from Bat-CoV. The most likely scenario is that a progenitor of SARS-CoV-2 jumped to humans from bats, possibly through pangolins. The virus then mutated and evolved in humans, till such time that it became capable of efficient transmission, and started causing widespread respiratory disease.

Was SARS-CoV-2 developed in a lab? If one tried to develop a highly pathogenic

Box 1. Protection from SARS-CoV-2 infection:

There are three important guidelines to protect from infection.

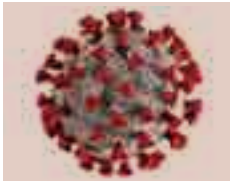
1. **Masks:** Though viruses are extremely small (nanometers) and are able to pass through very fine filters, SARS-CoV-2 and other respiratory viruses are released in large and small droplets while coughing, sneezing, speaking, or even breathing. Inhaling these droplets can infect a susceptible person. Wearing a mask, even a home-made cloth one, allows an infected person to contain the release of these droplets, and a susceptible person to prevent exposure. If both wear masks, the chances of transmission are negligible.
2. **Physical distancing:** Maintaining a distance of at least 2 m (or 6 feet) limits exposure to large droplets, loaded with the virus, from an infected person.
3. **Hand hygiene:** Virus-carrying droplets can contaminate surfaces such as door knobs, tabletops, etc. When a susceptible person touches these surfaces, the virus is transferred to their hands and, eventually, finds its way to their mouth or nose. Since SARS-CoV-2 has a lipid (fat) membrane, washing hands with soap and water destroys it just like soap cuts through grease. Hand sanitizers with about 70% alcohol are also helpful.

coronavirus, the logical starting point would be another human coronavirus, not an obscure bat or pangolin virus that has never been found to cause disease in humans. The Nobel laureate Prof. Luc Montagnier has referred to the presence of elements of both the HIV-1 retrovirus and the malarial parasite *Plasmodium falciparum* in the SARS-CoV-2 genome. These short sequences, as discovered by scientists in 2005, are required for genome replication by many coronaviruses. Researchers in India also pointed this out erroneously and retracted their results. There is no credible evidence that SARS-CoV-2 was developed in a lab. All available evidence seems to point to this virus being a product of natural evolution.

Parting thoughts

Can a bat virus jump into humans again? Absolutely. So can viruses from rodents, monkeys, and other mammals. We must, therefore, carry out routine surveillance of our ecosystems to recognize potential jumps well in advance. For example, India has 117 indigenous species of bats, but we know very little about the viruses they harbour. Killing bats or other animals is not the solution; they have valuable roles in our ecosystem. Surveillance and ensuring that their habitats (i.e., forests) are not destroyed will reduce animal-human contact, and reduce the risk of animal viruses getting into human populations. This is the only way to avoid future disease outbreaks.

Key takeaways



- SARS-CoV-2 gets its name from the fact that it is the second coronavirus known to cause Severe Acute Respiratory Syndrome in humans. Like other coronaviruses, SARS-CoV-2 is an RNA virus with a crown-like appearance when viewed under an electron microscope.
- While SARS-CoV-2 has an unusually large genome, its mutation rate is on average lower than other RNA viruses. As of now, only one strain of SARS-CoV-2 is circulating globally, and one mutation (called D614G) has spread faster than others.
- SARS-CoV-2 can stay infectious on a variety of surfaces for variable durations, but can be easily killed with soaps or disinfectants.
- There is no credible evidence that SARS-CoV-2 was created in a lab. All available evidence points to it being a product of natural evolution that originated in bats, and jumped to humans directly, or via pangolins.
- Future outbreaks of zoonoses from wild animals can be avoided by routine surveillance of ecosystems and ensuring that their (forest) habitats are not destroyed.

Note: Source of the image used in the background of the article title: https://phil.cdc.gov/Details.aspx?pid=23312. Credits: Alissa Eckert, MSMI & Dan Higgins, MAMS, created at the Centers for Disease Control and Prevention (CDC). License: Public Domain.

Shahid Jameel is a former Group Leader of Virology at The International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi. He is now the CEO of the DBT/Wellcome Trust India Alliance.