THE COVID LAB-LEAK HYPOTHESIS: WHAT SCIENTISTS DO AND DON’T KNOW

***Nature Journal** examines arguments that the coronavirus SARS-CoV-2 escaped from a lab in China, and the science behind them.*

Debate over the idea that the coronavirus SARS-CoV-2 emerged from a laboratory has escalated over the past few weeks, coinciding with the annual World Health Assembly, at which the World Health Organization (WHO) and officials from nearly 200 countries discussed the COVID-19 pandemic. After last year’s assembly, the WHO agreed to sponsor the first phase of an investigation into the pandemic’s origins; this probe took place in China in early 2021. Most scientists say SARS-CoV-2 probably has a natural origin, and was transmitted from an animal to humans. However, a lab leak has not been ruled out, and many are calling for a deeper investigation into the hypothesis that the virus emerged from the Wuhan Institute of Virology (WIV), located in the Chinese city where the first COVID-19 cases were reported. On 26 May, US President Joe Biden tasked the US Intelligence Community to join efforts to find SARS-CoV-2’s origins, whatever they might be, and report back in 90 days.

Australia, the European Union and Japan have also called for a robust investigation into SARS-CoV-2’s origins. The WHO has yet to reveal the next phase of its investigation. But China has asked that the probe examine other countries. Such reticence, and the fact that China has withheld information in the past, has fuelled suspicions about a ‘lab leak’. For instance, Chinese government officials suppressed crucial public-health data at the start of the COVID-19 pandemic, and during the 2002–04 severe acute respiratory syndrome (SARS) epidemic, according to high-level reports.

Nature looks at the key arguments that support a lab leak, and the extent to which research has answers.

There’s not yet any substantial evidence for a lab leak. Why are scientists still considering it?

Scientists don’t have enough evidence about the origins of SARS-CoV-2 to rule out the lab-leak hypothesis, or to prove the alternative — that the virus has a natural origin. Many infectious-disease researchers agree that the most probable scenario is that the virus evolved naturally and spread from a bat either directly to a person or through an intermediate animal. Most emerging infectious diseases begin with a spillover from nature, as was seen with HIV, influenza epidemics, Ebola outbreaks and the coronaviruses that caused the SARS epidemic beginning in 2002 and the Middle East respiratory syndrome (MERS) outbreak beginning in 2012.

Researchers have some leads that support a natural origin. Bats are known carriers of coronaviruses, and scientists have determined that the genome of SARS-CoV-2 is most similar to that of RATG13, a coronavirus that was first found in a horseshoe bat (*Rhinolophus affinis*) in the southern Chinese province of Yunnan in 2013. But RATG13’s genome is only 96% identical to SARS-CoV-2’s, suggesting that a closer relative of the virus — the one passed to humans — remains unknown.

Still, the possibility remains that SARS-CoV-2 escaped from a lab. Although lab leaks have never caused an epidemic, they have resulted in small outbreaks involving well-documented viruses. A relevant example happened in 2004, when two researchers were independently infected by the virus that causes SARS at a virology lab in Beijing that studied the disease. They spread the infection to seven others before the outbreak was contained.

What are the key arguments for a lab leak?

In theory, COVID-19 could have come from a lab in a few ways. Researchers might have collected SARS-CoV-2 from an animal and maintained it in their lab to study, or they might have created it by engineering coronavirus genomes. In these scenarios, a person in the lab might have then been accidentally or deliberately infected with the virus, and then spread it to others — sparking the pandemic. There is currently no clear evidence to back these scenarios, but they aren’t impossible.

People have made a number of arguments for a lab origin for SARS-CoV-2 that are currently conjecture.

One holds that it’s suspicious that, almost a year and a half into the pandemic, SARS-CoV-2’s closest relative still hasn’t been found in an animal. Another suggests it is no coincidence that COVID-19 was first detected in Wuhan, the site of a top lab studying coronaviruses — the WIV.

Some lab-leak proponents contend that the virus contains unusual features and genetic sequences signalling that it was engineered by humans. And some say that SARS-CoV-2 spreads among people so readily that it must have been created with that intention. Another argument suggests that SARS-CoV-2 might be derived from coronaviruses found in an unused mine where WIV researchers collected samples from bats between 2012 and 2015.

So what do infectious-disease researchers and evolutionary biologists say about these arguments?

Is it suspicious that no animal has been identified as transmitting the virus to humans?

Outbreak-origin investigations often take years, and some culprits remain unknown. It took 14 years to nail down the origin of the SARS epidemic, which began with a virus in bats that spread to humans, most likely through civets. To date, a complete Ebola virus has never been isolated from an animal in the region where the world’s largest outbreak occurred between 2013 and 2016.

Origin investigations are complicated because outbreaks among animals that aren’t the main hosts of a particular virus, such as civets in the case of SARS, are often sporadic. Researchers must find the right animal before it dies or clears the infection. And, even if the animal tests positive, viruses found in saliva, faeces or blood are often degraded, making it difficult to sequence the pathogen’s whole genome.

Scientists have made some progress since the pandemic began, however. A report posted to the preprint server bioRxiv on 27 May suggests that RmYN02, a coronavirus in bats in southern China, might be more closely related to SARS-CoV-2 than RATG13 is.

As for finding an intermediate host animal, researchers in China have tested more than 80,000 wild and domesticated animals; none has been positive for SARS-CoV-2. But this number is a tiny fraction of the animals in the country. To narrow the search down, researchers say, more strategic testing is needed to isolate, for example, animals that come in close contact with people.

Is it suspicious that the WIV is in Wuhan?

Virology labs tend to specialize in the viruses around them, says Vincent Munster, a virologist at the Rocky Mountain Laboratories, a division of the US National Institutes of Health, in Hamilton, Montana. The WIV specializes in coronaviruses because many have been found in and around China. Munster names other labs that focus on endemic viral diseases: haemorrhagic fever labs in Africa and dengue-fever labs in Latin America, for example. “Nine out of ten times, when there’s a new outbreak, you’ll find a lab that will be working on these kinds of viruses nearby,” says Munster.

Researchers note that a coronavirus outbreak in Wuhan isn’t surprising, because it’s a city of 11 million people in a broader region where coronaviruses have been found. It contains an airport, train stations and markets selling goods and wildlife transported there from around the region — meaning a virus could enter the city and spread rapidly.

Does the virus have features that suggest it was created in a lab?

Several researchers have looked into whether features of SARS-CoV-2 signal that it was bioengineered. One of the first teams to do so, led by Kristian Andersen, a virologist at Scripps Research in La Jolla, California, determined that this was “improbable” for a few reasons, including a lack of signatures of genetic manipulation. Since then, others have asked whether the virus’s furin cleavage site is evidence of engineering, because SARS-CoV-2 has these sites but its closest relatives don’t. The furin cleavage site is important because it’s in the virus’s spike protein, and cleavage of the protein at that site is necessary for the virus to infect cells.

But there are furin cleavage sites in many other coronaviruses, such as those that cause colds. Viruses containing the site are scattered across the coronavirus family tree, rather than confined to a group of closely related viruses. Stephen Goldstein, a virologist at the University of Utah in Salt Lake City, says the site probably evolved multiple times because it provides an evolutionary advantage.

Another feature of SARS-CoV-2 that has drawn attention is a combination of nucleotides that underlie a segment of the furin cleavage site: CGG (these encode the amino acid arginine). A Medium article that speculates on a lab origin for SARS-CoV-2 (see go.nature.com/3xko) quotes David Baltimore, a Nobel laureate and professor emeritus of biology at the California Institute of Technology in Pasadena, as saying that viruses don’t usually have that particular code for arginine, but humans often do — a “smoking gun”, hinting that researchers might have tampered with SARS-CoV-2’s genome.

Andersen says that Baltimore was incorrect about that detail, however. In SARS-CoV-2, about 3% of the nucleotides encoding arginine are CGG, he says. And he points out that around 5% of those encoding arginine in the virus that caused the original SARS epidemic are CGG, too. In an e-mail to Nature, Baltimore says Andersen could be correct that evolution produced SARS-CoV-2, but adds that “there are other possibilities and they need careful consideration, which is all I meant to be saying”.

…