Can we know for sure COVID's origins? Why is Omicron so persistent? Knowing how evolution works provides guidance

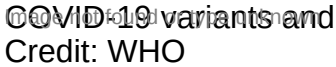
The latest phrase borrowed from biology in COVID conversations is *convergent evolution*. It refers to pairs of unrelated species that look similar because their ancestors evolved under similar environmental conditions. Natural selection favored adaptive (helpful) inherited traits, and millennia later, two unrelated species of mammals or birds look remarkably alike.

Convergent evolution happens to viruses, too. It is unspooling right now as SARS-CoV-2 genome evolution coalesces into variations on the Omicron theme.

The natural history of SARS-CoV-2 began with the wild type, another term from classical genetics. It means "most common," not "normal" as the media often misuses it.

As the virus changed, we grouped sets of new mutations, which substitute one RNA base of the genome at a time, into "variants." We named them, which biologists tend to do.

Alpha, recognized in November 2020, begat beta, gamma, and delta, all of which stayed with us for a bit. The next few versions were fleeting. The International Committee on Taxonomy of Viruses and WHO skipped Nu (because it sounds like "new") and Xi (a common surname), landing on Omicron. And natural selection has favored its collection of mutations. No new Greek letters necessary.


COVID-19 variants and their symptoms
Credit: WHO

When species look alike

Biologists term traits that are alike in two species that arise from recent shared ancestors *homologous*, while similar structures or behaviors that arise from similar environmental exposures are *analogous*. Convergent evolution reflects responses to similar environments (analogy), rather than descent from recent shared ancestors (homology).

Striking examples of convergent evolution are pairs of placental mammals and Australian marsupials. These include anteaters, moles, wolves, ocelots and native cats, flying squirrels and flying phalangers, and groundhogs and wombats.

Follow the latest news and policy debates on sustainable agriculture, biomedicine, and other 'disruptive' innovations. Subscribe to our newsletter.
SIGN UP

Species that illustrate convergent evolution need not be a world apart. Familiar examples abound.

- A shark is a fish and a dolphin a mammal, each adapted to life in the ocean.
- Butterflies, moths, hummingbirds, and the Australian opossum extract nectar from flowers using long

beak-like structures.
• Hedgehogs, porcupines, and echidnas are covered in quills.
• Bats and whales detect high frequencies of sound for echolocation.
• Pig snouts resemble those of peccaries.

Evolution is *ongoing* due to the flexibility of genetic material – RNA or DNA. Sequences of the building blocks change, unprovoked, when errors occur during the molecules' replication, like a glitch in copying and pasting a document. Mutations also arise from exposure to certain chemicals or radiation. Natural selection then retains altered genetic information that is helpful (adaptive) through capacity to reproduce, and weeds out individuals with detrimental genetic changes.

Evolution is also *branching* – it always is, the popular meme of apes morphing into humans a cringeworthy error. When a new mutation happens as a virus (or cell) replicates its genetic material, the old version remains in half the offspring. That's why even as the latest version of SARS-CoV-2 begins to circulate, its older brethren are still around. If a newbie has a reproductive advantage – spreads more readily from host to host – it eventually dominates the population of viruses in an area.

The origin question: We will never know

We still do not, and may never, know the immediate predecessor of SARS-CoV-2. I suspect many other geneticists share my opinion that the origin is unfathomable, but have been loathe to speak out because of politics and a public uncomfortable with the reality that science is inherently uncertain.

So here's what we *do* know. The answer may lie in a very specific part of the enemy.

SARS-CoV-2 presumably descended from a different coronavirus, possibly in a cave or caves. One candidate is RaTG13, but there are likely others. We can't describe what we haven't yet discovered.

Clues to the origin lie in what to the virus is the most important part of its anatomy: the triplets of spike proteins proteins that festoon the surface. Each spike has two parts: one grabs a cell, then the other slips the virus inside.

A spike grabs on using its receptor binding domain (RBD), which adheres to receptors (ACE2) on many host cell types, like in our lungs. A small part of the RBD, the "furin cleavage site," holds the halves of a spike together, until a host's enzyme cuts it. That jettisons the virus into the cell like an arrow shot from a bow.

Key to the power of the furin cleavage site is a specific sequence of four contiguous amino acids. This is where looking at other viruses provides compelling clues to where SARS-CoV-2 may have come from.

The first SARS virus, from 2003, doesn't have a furin cleavage site, but HIV, the Ebola virus, and the MERS coronavirus do. That's why the furin cleavage site in SARS-CoV-2 is dubbed a *"gain-of-function."* Altering this viral Achilles heel can and has changed the speed of transmission.

sars-cov-2 anatomy

Image not found or type unknown
Credit: Global Health News Wire

(ASIDE: Like convergent evolution, gain-of-function isn't a new term. In grad school in the late 1970s I made gain-of-function mutations by bombarding flies with X-rays. Now Etsy sells a gain-of-function tee shirt that idiotically besmirches Anthony Fauci.)

Back to the million dollar question: How did the novel furin cleavage site get into SARS-CoV-2, when its nearest known relative doesn't have it? Perhaps more importantly, *where* did this happen?

Did pieces of RNA altering the furin cleavage site swap into the genomes of the recent ancestors of SARS-CoV-2:

• in a bat cave?
• in an abandoned mine shaft in Yunnan province?
• in a "wet market" near the Wuhan Institute of Virology?
• directly in a lab at the institute, either by accident or purposely?

More than one of the above, none of the above, or even all of the above?

However it happened, most puzzling is the extent of change to the furin cleavage site: *four contiguous amino acids.*

That represents twelve RNA bases, because RNA or DNA triplets encode the amino acids that make up proteins.

Did the four-amino-acid-change happen one-at-a-time, and we just haven't identified the intermediates? Or all at once? Stretches of genetic material that move among genomes are well known for organisms, from bacteria to humans. But they are rare among viruses.

So how did the exact stretch of amino acids that catapult the viruses into our cells get there?

Occum's razor might suggest that someone engineered the four amino acids into the furin cleavage site, perhaps by accident, perhaps to create a bioweapon. Then a "lab leak" transpired.

But because we can't discount a role for natural selection acting on a protein in a way that benefits the virus once it got into the environment, we can't know with any degree of confidence how, exactly, SARS-CoV-2 came to attack humanity and many other species. That's the crux of the issue that I rarely read in media reports.

From origin to Omicron. What's next?

However SARS-CoV-2 got here, it has gone through a complex choreography of changes.

"G614D" was an early mutant slipping unnoticed into New York City from Europe as politicians unschooled in biology frantically closed borders to flights from certain nations, as if that could keep viruses out like closing the doors on a plane. ("G614D" is code for the particular amino acid change.)

The virus continued to reinvent itself. We named mutations, grouped them into variants, and tracked their spread around the world through their genome sequences on increasingly crowded dashboards, the US lagging behind for uncomfortable months. SARS-CoV-2 variants soon began marching up the Greek alphabet towards Omicron, 15th of the 24 letters.

Then the pace of mutation slowed as natural selection continued to favor changes that propelled the virus into and out of human noses and mouths ever faster, the same changes persisting among even the least related viral strains, because they're adaptive. The retooled viruses began to make people less sick, while vaccines perhaps dampened the spread and tamed symptoms.

And so after three years of change, SARS-CoV-2 evolution has slowed as a suite of similar Omicrons persist, because the comprising mutations enhance viral binding to our ACE2 receptors. *If it ain't broke …*

Looking at the virus as a whole, the extent of change in a short time is astounding, although perhaps that's partly because we haven't given such genetic scrutiny to other pathogens, posting new genome sequences daily.

For SARS-CoV-2, all possible single-RNA-base changes that could have altered the amino acid sequence of just the viral spike protein had happened by early 2021. But the beast hasn't ceased. Mutations that alter *two* RNA bases at a time are now accelerating. They have *"gained momentum and their numbers are increasing rapidly. These provide a large mutation landscape for SARS-CoV-2 future evolution, on which research should focus now,"*

write Jiri Zahradnik, Jaroslav Nunvar, and Gideon Schreiber in Perspectives: SARS-CoV-2 Spike Convergent Evolution as a Guide to Explore Adaptive Advantage.

In other words, the virus isn't yet done with us. Anthony Fauci's agrees, in his farewell Perspective in the New England Journal of Medicine, which borrows from Yogi Berra: *"It Ain't Over Till It's Over…but It's Never Over — Emerging and Reemerging Infectious Diseases."* Dr. Fauci calls COVID *"the loudest wake-up call in more than a century to our vulnerability to outbreaks of emerging infectious diseases."*

As long as we share the planet with others and evolution continues to mold genomes, infectious diseases will be with us. But we control that future, to an extent. Writes Dr. Fauci:

"The emergence of new infections and the reemergence of old ones are largely the result of human interactions with and encroachment on nature. As human societies expand in a progressively interconnected world and the human–animal interface is perturbed, opportunities are created, often aided by climate changes, for unstable infectious agents to emerge, jump species, and in some cases adapt to spread among humans. When it comes to emerging infectious diseases, it's never over."

I agree.

Ricki Lewis has a PhD in genetics and is a genetics counselor, science writer and author of Human Genetics: The Basics. Follow her at her website or Twitter @rickilewis

A version of this article appeared originally at PLOS and is posted here with permission. Check out PLOS on Twitter @PLOS