

The spike protein of SARS-CoV-2 has a common mutation (circled) that seems to shift the protein from a closed (left) to an open (right) form.

SOURCE: STRUCTURAL DATA FROM K. SHEN & J. LUBAN

# MAKING SENSE OF CORONAVIRUS MUTATIONS

Different SARS-CoV-2 strains haven't yet had a major impact on the course of the pandemic – but they might in future. **By Ewen Callaway**

**W**hen COVID-19 spread around the globe this year, David Montefiori wondered how the deadly virus behind the pandemic might be changing as it passed from person to person. Montefiori is a virologist who has spent much of his career studying how chance mutations in HIV help it to evade the immune system. The same thing might happen with SARS-CoV-2, he thought.

In March, Montefiori, who directs an AIDS-vaccine research laboratory at Duke University in Durham, North Carolina, contacted Bette Korber, an expert in HIV evolution and a long-time collaborator. Korber, a computational biologist at the Los Alamos National Laboratory (LANL) in Sante Fe, New Mexico,

had already started scouring thousands of coronavirus genetic sequences for mutations that might have changed the virus's properties as it made its way around the world.

Compared with HIV, SARS-CoV-2 is changing much more slowly as it spreads. But one mutation stood out to Korber. It was in the gene encoding the spike protein, which helps virus particles to penetrate cells. Korber saw the mutation appearing again and again in samples from people with COVID-19. At the 614th amino-acid position of the spike protein, the amino acid aspartate (D, in biochemical shorthand) was regularly being replaced by glycine (G) because of a copying fault that altered a single nucleotide in the virus's 29,903-letter RNA code. Virologists were calling it the D614G mutation.

In April, Korber, Montefiori and others warned in a preprint posted to the bioRxiv server that "D614G is increasing in frequency at an alarming rate"<sup>1</sup>. It had rapidly become the dominant SARS-CoV-2 lineage in Europe and had then taken hold in the United States, Canada and Australia. D614G represented a "more transmissible form of SARS-CoV-2", the paper declared, one that had emerged as a product of natural selection.

These assertions dismayed many scientists. It wasn't clear that the D614G viral lineage was more transmissible, or that its rise indicated anything unusual, they said. But alarm spread fast across the media. Although many news stories included researchers' caveats, some headlines declared that the virus was mutating to become more dangerous. In retrospect,

Montefiori says he and his colleagues regret describing the variant's rise as "alarming". The word was scrubbed from the peer-reviewed version of the paper, published in *Cell* in July<sup>2</sup>.

The work sparked a frenzy of interest in D614G. Even those who were sceptical that the mutation had changed the virus's properties agreed that it was intriguing, because of its meteoric rise and ubiquity. For months, that lineage has been found in almost all sequenced samples of SARS-CoV-2 (see 'Global spread'). "This variant now is the pandemic. As a result, its properties matter," wrote Nathan Grubaugh, a viral epidemiologist at the Yale School of Public Health in New Haven, Connecticut, and two colleagues in a *Cell* essay on Korber and Montefiori's findings<sup>3</sup>.

So far, the upshot of this work is less clear than Montefiori and Korber's preprint suggested. Some experiments suggest that viruses carrying the variant infect cells more easily. Other work has revealed possible good news: the variant might mean that vaccines can target SARS-CoV-2 more easily. But many scientists say there remains no solid proof that D614G has a significant effect on the spread of the virus, or that a process of natural selection explains its rise. "The jury's out," says Timothy Sheahan, a coronavirologist at the University of North Carolina at Chapel Hill. "This mutation might mean something, or it might not."

Researchers still have more questions than answers about coronavirus mutations, and no one has yet found any change in SARS-CoV-2 that should raise public-health concerns, Sheahan, Grubaugh and others say. But studying mutations in detail could be important for controlling the pandemic. It might also help to pre-empt the most worrying of mutations: those that could help the virus to evade immune systems, vaccines or antibody therapies.

### Slow change

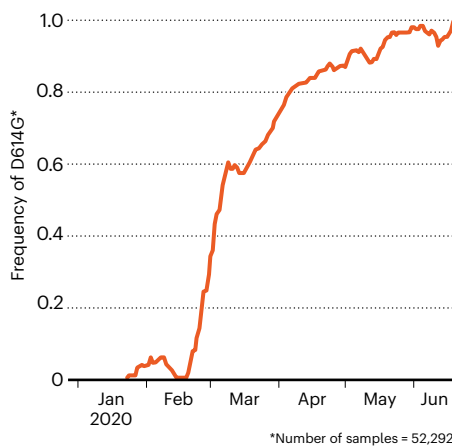
Soon after SARS-CoV-2 was detected in China, researchers began analysing viral samples and posting the genetic codes online. Mutations – most of them single-letter alterations between viruses from different people – allowed researchers to track the spread by linking closely related viruses, and to estimate when SARS-CoV-2 started infecting humans.

Viruses that encode their genome in RNA, such as SARS-CoV-2, HIV and influenza, tend to pick up mutations quickly as they are copied inside their hosts, because enzymes that copy RNA are prone to making errors. After the severe acute respiratory syndrome (SARS) virus began circulating in humans, for instance, it developed a kind of mutation called a deletion that might have slowed its spread<sup>4</sup>.

But sequencing data suggest that coronaviruses change more slowly than most other RNA viruses, probably because of a 'proofreading' enzyme that corrects potentially fatal copying

### GLOBAL SPREAD

By the end of June, the D614G mutation was found in almost all SARS-CoV-2 samples worldwide.



mistakes. A typical SARS-CoV-2 virus accumulates only two single-letter mutations per month in its genome – a rate of change about half that of influenza and one-quarter that of HIV, says Emma Hodcroft, a molecular epidemiologist at the University of Basel, Switzerland.

Other genome data have emphasized this stability – more than 90,000 isolates have been sequenced and made public (see [www.gisaid.org](http://www.gisaid.org)). Two SARS-CoV-2 viruses collected from anywhere in the world differ by an average of just 10 RNA letters out of 29,903, says Lucy Van Dorp, a computational geneticist at University College London, who is tracking the differences for signs that they confer an evolutionary advantage.

Despite the virus's sluggish mutation rate, researchers have catalogued more than 12,000 mutations in SARS-CoV-2 genomes. But scientists can spot mutations faster than they can

**"This variant now is the pandemic. As a result, its properties matter."**

make sense of them. Many mutations will have no consequence for the virus's ability to spread or cause disease, because they do not alter the shape of a protein, whereas those mutations that do change proteins are more likely to harm the virus than improve it (see 'A catalogue of coronavirus mutations'). "It's much easier to break something than it is to fix it," says Hodcroft, who is part of Nextstrain (<https://nextstrain.org>), an effort to analyse SARS-CoV-2 genomes in real time.

Many researchers suspect that if a mutation did help the virus to spread faster, it probably happened earlier, when the virus first jumped into humans or acquired the ability to move efficiently from one person to another. At a time when nearly everyone on the planet is susceptible, there is likely to be little evolutionary pressure on the virus to spread better, so

even potentially beneficial mutations might not flourish. "As far as the virus is concerned, every single person that it comes to is a good piece of meat," says William Hanage, an epidemiologist at the Harvard T. H. Chan School of Public Health in Boston, Massachusetts. "There's no selection to be doing it any better."

### Faster spread?

When Korber saw the rapid spread of D614G, she thought she might have found an example of meaningful natural selection. The mutation caught her eye because of its position in the spike protein, which is a major target for 'neutralizing' antibodies that bind to the virus and render it non-infectious. And viruses with the mutation were also rising in frequency in more than one part of the world.

D614G was first spotted in viruses collected in China and Germany in late January; most scientists suspect the mutation arose in China. It's now almost always accompanied by three mutations in other parts of the SARS-CoV-2 genome – possible evidence that most D614G viruses share a common ancestor.

D614G's rapid rise in Europe drew Korber's attention. Before March – when much of the continent went into lockdown – both unmutated 'D' viruses and mutated 'G' viruses were present, with D viruses prevalent in most of the western European countries that geneticists sampled at the time. In March, G viruses rose in frequency across the continent, and by April they were dominant, reported Korber, Montefiori and their team<sup>1,2</sup>.

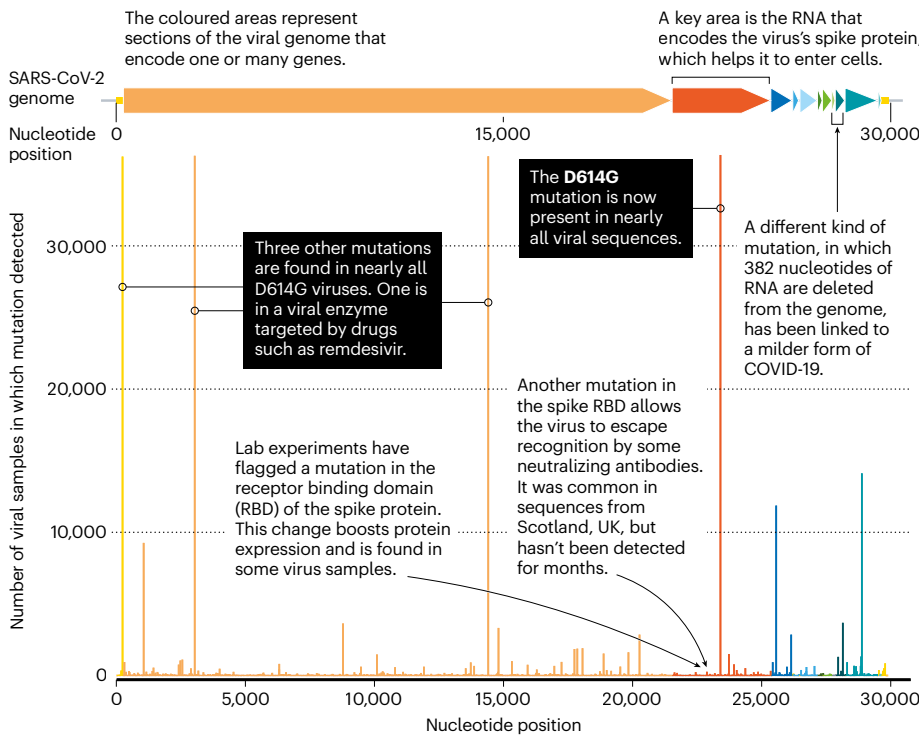
But natural selection in favour of G viruses isn't the only, or even the most likely, explanation for this pattern. The European dominance of G variants could be simply down to chance – if, for instance, the mutation happened to be slightly more common in the viruses that arrived in Europe. A small number of individuals seem to be responsible for most of the virus's spread, and an early, chance tilt in favour of G viruses could explain the lineage's apparent takeover now. Such 'founder effects' are common in viruses, especially when they spread unchecked, as SARS-CoV-2 did in much of Europe until mid- to late March.

Korber and her colleagues tried to rule out a founder effect, by showing in their April preprint<sup>1</sup> that D614G rose to dominance quickly in Canada, Australia and parts of the United States (an exception was Iceland, where G viruses present early in its outbreak were overtaken by D viruses). Analysing hospitalization data from Sheffield, UK, the team found no evidence that viruses carrying the mutation made people any sicker. But those infected with G viruses seemed to have slightly higher levels of viral RNA in their noses and mouths than did those with D viruses.

Many scientists weren't convinced that D614G's rise was remarkable – or all that relevant to the pandemic. "I thought that preprint

## A CATALOGUE OF CORONAVIRUS MUTATIONS

Various mutations have been detected in SARS-CoV-2 genomes, including the most prevalent one, D614G. The virus's genetic code has just under 30,000 nucleotides of RNA, or letters, that spell out at least 29 genes. The most common mutations are single-nucleotide changes.



was incredibly premature,” says Sheahan.

Montefiori says his and Korber’s perspective on D614G was shaped by their work on HIV, which has found that even seemingly insignificant mutations can have a profound effect on how the immune system recognizes that virus. “We were alarmed by it, and we need to see if it’s having an effect on vaccines,” he says.

### Rush of lab studies

To examine further whether D614G made the virus more transmissible, Montefiori gauged its effects under laboratory conditions. He couldn’t study the natural SARS-CoV-2 virus in his lab, because of the biosafety containment required. So he studied a genetically modified form of HIV that used the SARS-CoV-2 spike protein to infect cells. Such ‘pseudovirus’ particles are a workhorse of virology labs: they enable the safe study of deadly pathogens such as the Ebola virus, and they make it easy to test the effects of mutations.

The first team to report pseudovirus experiments on D614G, in June, was led by Hyeryun Choe and Michael Farzan, virologists at the Scripps Research Institute in La Jolla, California<sup>5</sup>. Several other teams have posted similar studies on bioRxiv (Montefiori’s experiments, and those of another collaborator, appeared in the *Cell* paper<sup>2</sup>). The teams used different pseudovirus systems and tested them on various kinds of cell, but the experiments pointed to the same conclusion: viruses carrying the G mutation infected cells much more ably than did D viruses – up to ten times more efficiently, in some cases.

In laboratory tests, “all of us agree that D to G is making the particles more infectious”, says Jeremy Luban, a virologist at the University of Massachusetts Medical School in Worcester. But these studies come with many caveats – and their relevance to human infections is unclear. “What’s irritating are people taking their results in very controlled settings, and saying this means something for the pandemic. That, we are so far away from knowing,” says Grubaugh. The pseudoviruses carry only the coronavirus spike protein, in most cases, and so the experiments measure only the ability of these particles to enter cells, not aspects of their effects inside cells, let alone on an organism. They also lack the other three mutations that almost all D614G viruses carry. “The bottom line is, they’re not the virus,” says Luban.

Some labs are now working with infectious SARS-CoV-2 viruses that differ by only the single amino acid. These are tested in laboratory cultures of human lung and airway cells, and in lab animals such as ferrets and hamsters. For labs with the experience and the biosafety capabilities to manipulate viruses, “this is like bread-and-butter kind of work”, says Sheahan. The first of those studies, led by researchers at the University of Texas Medical Branch in Galveston, was reported in a 2 September preprint<sup>6</sup>. It found that viruses with the mutation were more infectious than were D viruses in a human lung cell line and in airway tissues, and that mutated viruses were present at greater levels in the upper airways of infected hamsters<sup>6</sup>.

Even these experiments might not offer

absolute clarity. Some studies show that certain mutations to the spike protein in the Middle East respiratory syndrome (MERS) virus can cause more-severe disease in mice – yet other mutations in the protein show very little effect in people or in camels, the likely reservoir for human MERS infections, says Stanley Perlman, a coronavirologist at the University of Iowa in Iowa City.

The clearest sign that D614G has an effect on the spread of SARS-CoV-2 in humans comes from an ambitious UK effort called the COVID-19 Genomics UK Consortium, which has analysed genomes of around 25,000 viral samples. From these data, researchers have identified more than 1,300 instances in which a virus entered the United Kingdom and spread, including examples of D- and G-type viruses.

A team led by Andrew Rambaut, an evolutionary biologist at the University of Edinburgh, UK, epidemiologist Erik Volz, at Imperial College London, and biologist Thomas Connor at Cardiff University, studied the UK spread of 62 COVID-19 clusters seeded by D viruses and 245 by G viruses<sup>7</sup>. The researchers found no clinical differences in people infected with either virus. However, G viruses tended to transmit slightly faster than lineages that didn’t carry the change, and formed larger clusters of infections. Their estimates of the difference in transmission rates hover around 20%, Volz says, but the true value could be a bit higher or lower. “There’s not a large effect in absolute terms,” says Rambaut.

It’s possible that D614G is an adaptation that helps the virus to infect cells or compete with viruses that don’t carry the change, while altering little about how SARS-CoV-2 spreads between people or through a population, Rambaut says. “This might be a bona fide adaptation to humans or some human cells,” agrees Grubaugh, “but that doesn’t mean anything changes. An adaptation doesn’t have to make it more transmissible.”

Grubaugh thinks that D614G has received too much attention from scientists, in part because of the high-profile papers it has garnered. “Scientists have this crazy fascination with these mutations,” he says. But he also sees D614G as a way to learn about a virus that doesn’t have much in the way of genetic diversity. “The virologist in me looks at these things and says it would be a lot of fun to study,” he says. “It creates this whole rabbit hole of different things you can go into.”

He’ll have company. Intense study of D614G should help to explain how SARS-CoV-2 fuses with cells, says Luban – a process that might be blocked by drugs or targeted by a vaccine. In an updated version of their pseudovirus experiments posted on bioRxiv on 16 July<sup>8</sup>, Luban’s team used cryo-electron microscopy to analyse the structure of spike proteins bearing the D614G change. The spike protein is comprised of three identical peptides in an ‘open’

SOURCES: L. VAN DORP ET AL. (HTTP://GO.NATURE.COM/3GSRNH6); REFS 2, 11, 12; B. E. YOUNG ET AL. LANCET 396, 603–611 (2020)

or 'closed' orientation. Previous research has suggested that at least two of the three peptides need to be open for the viral particle to fuse with the cell membrane<sup>9</sup>, and Luban's team found that viruses carrying the G spike variant were much more likely to be in this state (see 'The mutation that loosens the spike protein'). Computational modelling work by Montefiori and Korber, led by Korber's LANL colleague Sandrasegaram Gnanakaran, came to the same conclusion<sup>10</sup>. "It looks like this molecular machine is primed to go in a way that D is not," Luban says.

### No escape from antibodies – yet

Most available evidence suggests that D614G doesn't stop the immune system's neutralizing antibodies from recognizing SARS-CoV-2, as Montefiori had worried. That might be because the mutation is not in the spike protein's receptor-binding domain (RBD), a region that many neutralizing antibodies target: the RBD binds to the cell-receptor protein ACE2, a key step in the virus's entry to cells.

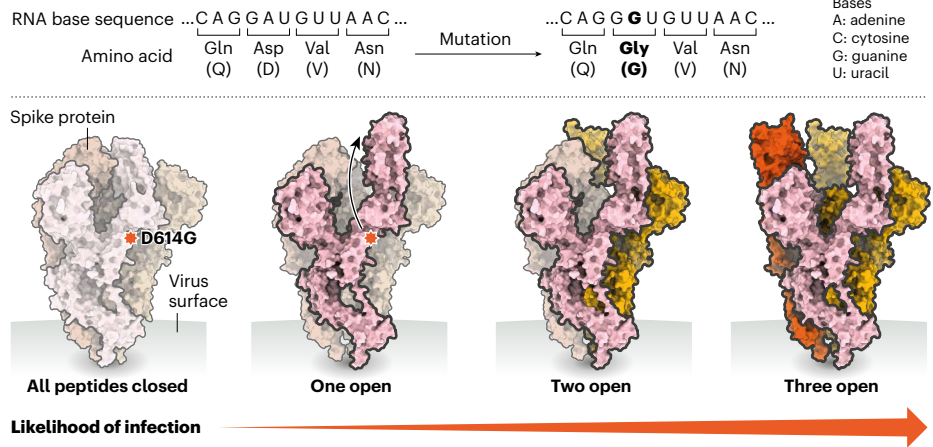
But evidence is emerging that other mutations could help the virus to avoid some antibodies. A team led by virologists Theodora Hatzioannou and Paul Bieniasz, at Rockefeller University in New York City, genetically modified the vesicular stomatitis virus – a livestock pathogen – so that it used the SARS-CoV-2 spike protein to infect cells, and grew it in the presence of neutralizing antibodies. Their goal was to select for mutations that enabled the spike protein to evade antibody recognition. The experiment generated spike-protein mutants that were resistant to antibodies taken from the blood of people who had recovered from COVID-19, as well as to potent 'monoclonal' antibodies that are being developed into therapies. Every one of the spike mutations was found in virus sequences isolated from patients, report Hatzioannou, Bieniasz and their team – although at very low frequencies that suggest positive selection is not yet making the mutations more common<sup>11</sup>.

Other scientists are trying to stay ahead of SARS-CoV-2's evolution by predicting which mutations are likely to be important. Jesse Bloom, an evolutionary virologist at the Fred Hutchinson Cancer Research Center in Seattle, Washington, led a team that created nearly 4,000 mutated versions of the spike protein's RBD, and measured how the alterations affected the expression of the spike protein and its ability to bind to ACE2. Most of the mutations had no effect on or hindered these properties, although a handful improved them<sup>12</sup>. Some of these mutations have been identified in people with COVID-19, but Bloom's team found no signs of natural selection for any of the variants. "Probably the virus binds to ACE2 about as well as it needs to right now," he says.

The researchers didn't test whether any of the mutations allow the virus to thwart the action of

## THE MUTATION THAT LOOSENS THE SPIKE PROTEIN

Spike proteins on SARS-CoV-2 bind to receptors on human cells, helping the virus to enter. A spike protein is made up of three smaller peptides in 'open' or 'closed' orientations; when more are open, it's easier for the protein to bind. The D614G mutation – the result of a single-letter change to the viral RNA code – seems to relax connections between peptides. This makes open conformations more likely and might increase the chance of infection.



antibodies, but his team's results suggest that such changes are possible. "It is a possibility, but by no means a certainty, that the virus will acquire mutations that change its susceptibility to antibodies and immunity," says Bloom.

Based on experience with other coronaviruses, that might take years. Studies of common-cold coronaviruses, sampled across multiple seasons, have identified some signs of evolution in response to immunity. But the pace of change is slow, says Volker Thiel, an RNA virologist at the Institute of Virology and Immunology in Bern. "These strains remain constant, more or less."

With most of the world still susceptible to SARS-CoV-2, it's unlikely that immunity is currently a major factor in the virus's evolution. But as population-wide immunity rises,

**"It is a possibility that the virus will acquire mutations that change its susceptibility to antibodies and immunity."**

whether through infection or vaccination, a steady trickle of immune-evading mutations could help SARS-CoV-2 to establish itself permanently, says Sheahan, potentially causing mostly mild symptoms when it infects individuals who have some residual immunity from a previous infection or vaccination. "I wouldn't be surprised if this virus is maintained as a more common, cold-causing coronavirus." But it's also possible that our immune responses to coronavirus infections, including to SARS-CoV-2, aren't strong or long-lived enough to generate selection pressure that leads to significantly altered virus strains.

Worrisome mutations could also become more common if antibody therapies aren't used wisely – if people with COVID-19 receive one antibody, which could be thwarted by a

single viral mutation, for example. Cocktails of monoclonal antibodies, each of which can recognize multiple regions of the spike protein, might lessen the odds that such a mutation will be favoured through natural selection, researchers say. Vaccines arouse less concern on this score because, like the body's natural immune response, they tend to elicit a range of antibodies.

It's even possible that the D614G change could make the virus an easier target for vaccines, Montefiori's team found in a study posted to bioRxiv in July<sup>13</sup>. Mice, monkeys and humans that received one of a number of experimental RNA vaccines, including one being developed by drug maker Pfizer in New York City, produced antibodies that proved more potent at blocking G viruses than D viruses.

With G viruses now ubiquitous, the finding is "good news", says Montefiori. But as a scientist who has watched HIV mutate to elude many vaccines developed against it, he remains wary of the potential of SARS-CoV-2 to evade humanity's responses. Luban agrees: "We need to keep our eyes open for additional changes."

Ewen Callaway writes for *Nature* from London.

- Korber, B. et al. Preprint at bioRxiv <https://doi.org/10.1101/2020.04.29.069054> (2020).
- Korber, B. et al. *Cell* **182**, 812–827 (2020).
- Grubaugh, N., Hanage, W. & Rasmussen, A. *Cell* **182**, 794–795 (2020).
- Muth, D. et al. *Sci. Rep.* **8**, 15177 (2018).
- Zhang, L. et al. Preprint at bioRxiv <https://doi.org/10.1101/2020.06.12.148726> (2020).
- Plante, K. et al. Preprint at bioRxiv <https://doi.org/10.1101/2020.09.01.278689> (2020).
- Volz, E. M. et al. Preprint at bioRxiv <https://doi.org/10.1101/2020.07.31.20166082> (2020).
- Yurkovetskiy, L. et al. Preprint at bioRxiv <https://doi.org/10.1101/2020.07.04.187757> (2020).
- Wrapp, D. et al. *Science* **367**, 1260–1263 (2020).
- Mansbach, R. A. et al. Preprint at bioRxiv <https://doi.org/10.1101/2020.07.26.219741> (2020).
- Weissblum, Y. et al. Preprint at bioRxiv <https://doi.org/10.1101/2020.07.21.214759> (2020).
- Starr, T. N. et al. *Cell* **182**, 1295–1310 (2020).
- Weissmann, D. et al. Preprint at bioRxiv <https://doi.org/10.1101/2020.07.22.20159905> (2020).