How SARS-CoV-2 first adapted in humans

Hyeryun Choe, Michael Farzan

+ See all authors and affiliations

Science 30 Apr 2021:
Vol. 372, Issue 6541, pp. 466-467
DOI: 10.1126/science.abi4711

Viruses need entry proteins to penetrate the cells where they will replicate. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) version is called the spike or S protein. The S protein, also the target of the current vaccines, is quickly adapting to its new human hosts. It took its first major step in this direction early in 2020, when its amino acid 614 (of 1297) changed from an aspartic acid (D) to a glycine (G). Viruses bearing this D614G mutation transmit among humans more rapidly and now form the majority in circulation. On page 525 of this issue, Zhang *et al.* (*1*) use careful structural analyses to reveal how D614G changed the S protein to accelerate the pandemic.

This change kept the S-protein S1 domain, which contains the RBD and binds ACE2, covalently linked to its S2 domain, which anchors the S protein to the virion. Notably, some—but not all—of these furin-site mutations significantly improved pseudovirus infection of cells (*4*).

This fix solved a technical problem, but it deepened a mystery. Although a number of distantly related coronaviruses carry furin cleavage sites at their S1-S2 boundaries, the SARS-CoV-1 S protein, and those of all known bat-derived viruses from the same Sarbecovirus lineage, lack this site. Instead of being cleaved in virus-producing cells, their S proteins are cleaved by different proteases while the virus is engaging ACE2 in the next, yet-to-be-infected cell (*5*). As it happened, furin-site mutations that improved SARS-CoV-2 S-protein function in pseudoviruses allowed the modified S protein to work with these later-stage enzymes, just like the SARS-CoV-1 version. Why then did the SARS-CoV-2 furin site persist, even though it made infection in cell culture less efficient? Indeed, viruses passaged in culture regularly lost this site. Does it somehow improve viral transmission? Would it eventually disappear over the course of the pandemic?

The underlying mechanism for this fitness advantage remained a point of controversy. Here, a second unusual property of the S protein, in this case shared with SARS-CoV-1, became relevant. The SARS-CoV-2 S protein, like most entry proteins of viruses with a lipid membrane, assembles into trimers. Typically, during the process of virion assembly, viral entry proteins subtly change their conformations, but it is unusual for these proteins to break their three-fold symmetry before they bind their receptor. However, the mature SARS-CoV-2 S protein often assumes an asymmetrical arrangement whereby one of its three RBDs assumes an open or "up" conformation (*1*, *9*). Only RBDs in this up conformation can bind ACE2. Once it does so, the S1 domains dissociate from S2, and S2 undergoes a pronounced rearrangement to a "postfusion" state. The released energy of this rearrangement drives the viral and cell membranes to fuse and gives the virus access to the cell interior.

Although SARS-CoV-2 shares high sequence homology with SARS-CoV, which caused the 2002–2004 SARS outbreak, the coronavirus family is diverse in both sequence and in host receptor preference. For example, SARS-CoV-2 and a "common cold" human coronavirus, HCoV-NL63, both recognize angiotensin-converting enzyme 2 (ACE2) as the host cell receptor, but SARS-CoV-2 and HCoV-NL63 belong to different coronavirus genera and have major sequence and structural differences in the receptor-binding domain (RBD) of S, sharing <30% sequence homology (2). This diversity in S indicates that coronaviruses have broad potential to tolerate changes in both sequence and structure without substantial loss of function. This may partially explain why coronaviruses can undergo zoonotic transmission and suggests that the full evolutionary potential of SARS-CoV-2 has yet to be revealed.

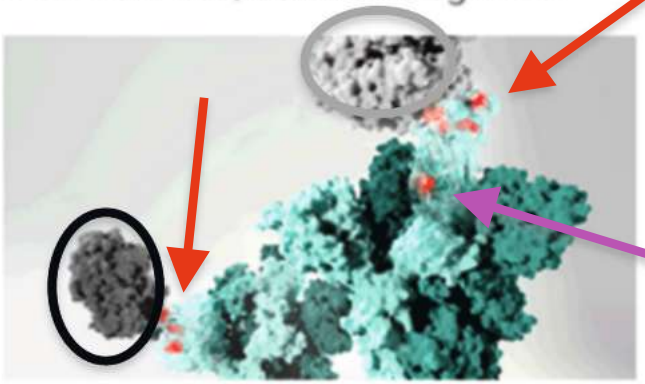
The S protein comprises two subunits: S1, which contains the RBD, and S2, which mediates virus–host cell fusion. Antibody-neutralizing epitopes are scattered throughout S but are mostly concentrated within the RBD. Despite the potential for plasticity, after nearly a year of spread (from December 2019) to >100 million people, there was limited evidence for evolution of SARS-CoV-2 S. The only notable evolutionary event was the D614G (Asp^{614}→Gly) substitution in S1, which increases ACE2 affinity, leading to higher infectivity and transmissibility. Viral sequences deposited in public databases were mostly obtained from the upper respiratory tract during acute infection, before major immune responses have occurred. Such sequences might not capture the effect of within-host immune selection on viral diversification.

The emerging plasticity of SARS-CoV-2

Kevin D. McCormick, Jana L. Jacobs, John W. Mellors

+ See all authors and affiliations

Science 26 Mar 2021:
Vol. 371, Issue 6536, pp. 1306-1308
DOI: 10.1126/science.abg4493



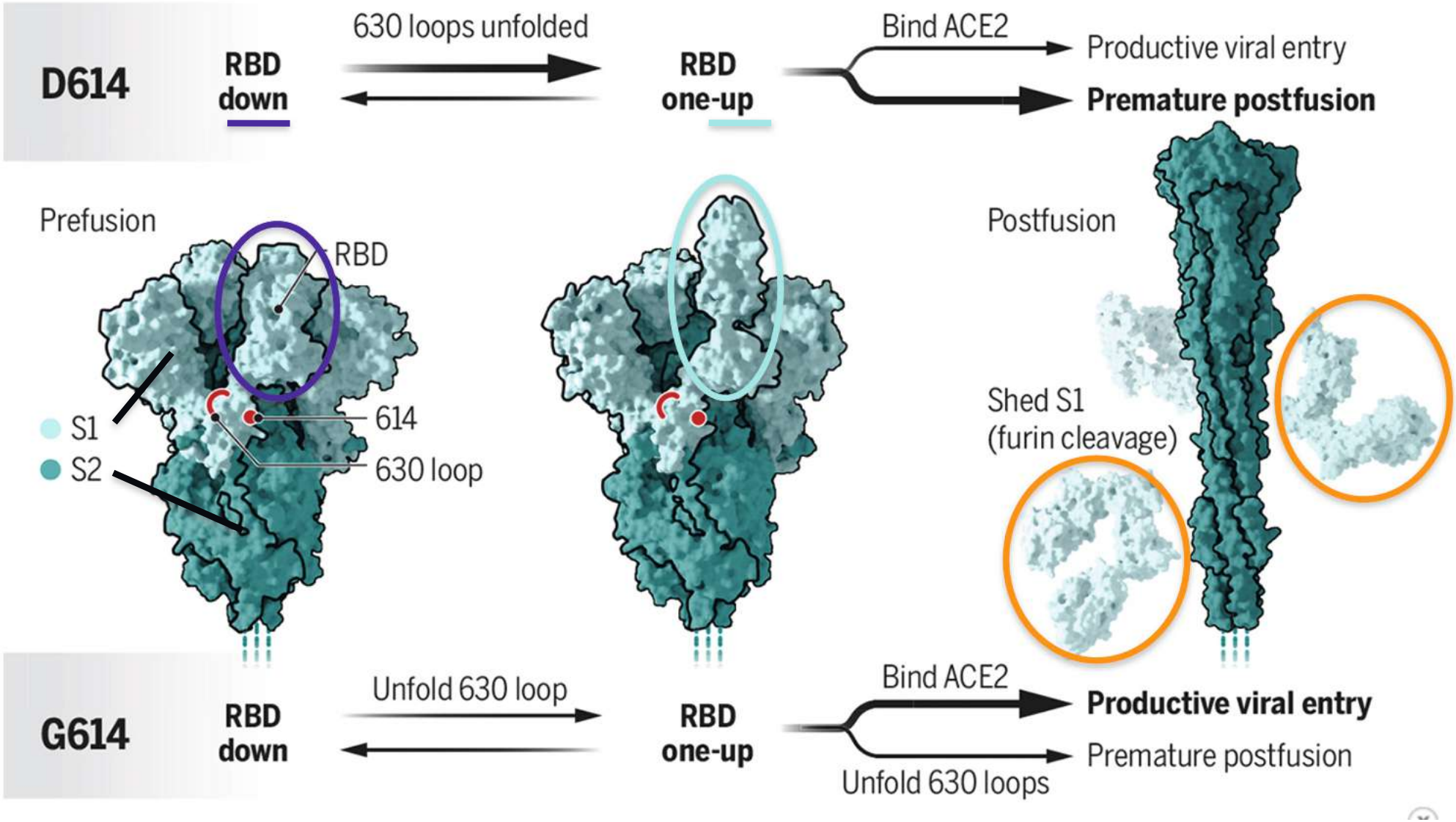
Mutations (red) in the spike protein (green) of SARS-CoV-2 variants that affect host receptor (light gray) or antibody (dark gray) binding could impair immunity.

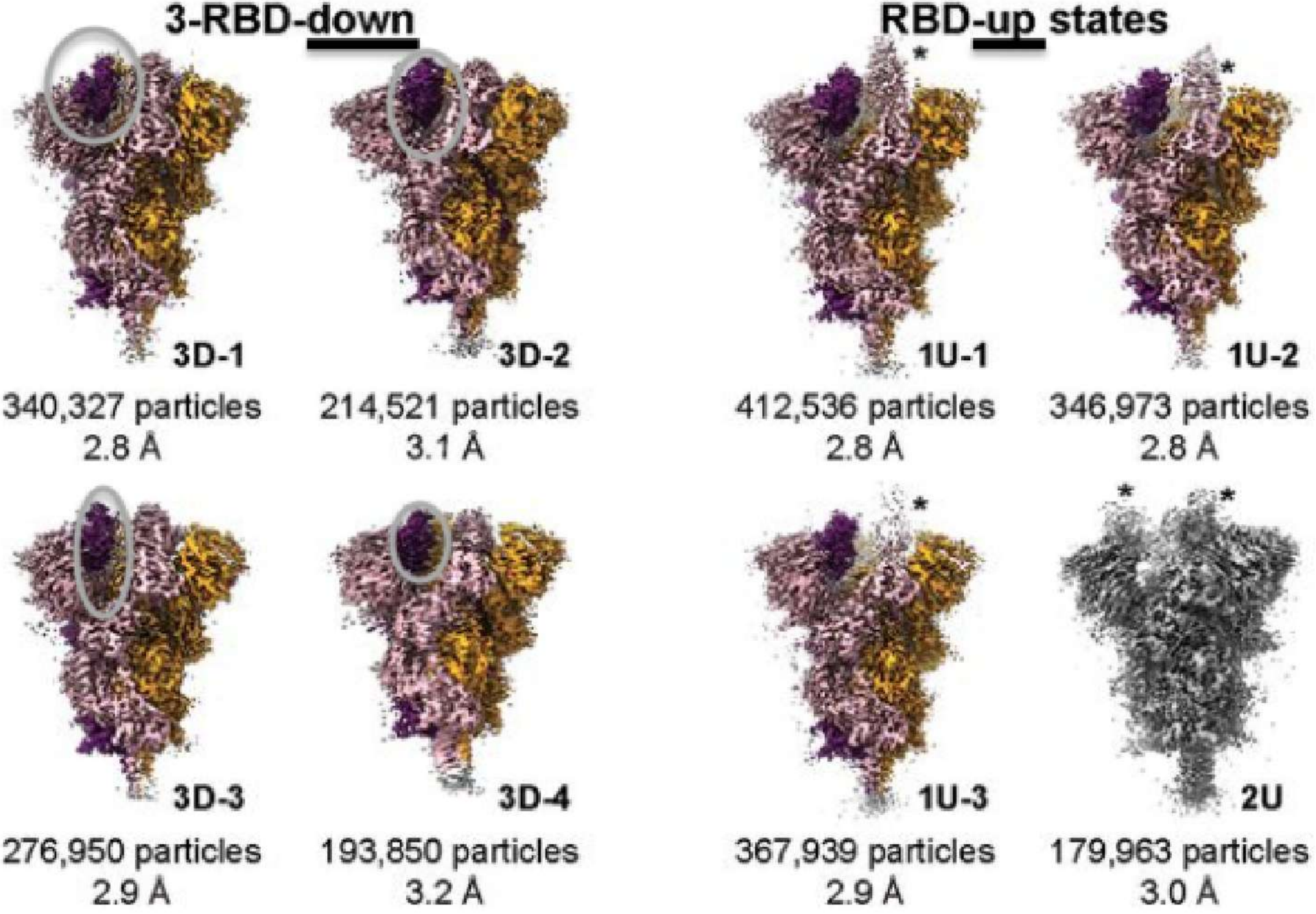
ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

Viruses evolve as a result of mutation (misincorporations, insertions or deletions, and recombination) and natural selection for favorable traits such as more efficient viral replication, transmission, and evasion of host defenses. Newly selected traits may be linked in unpredictable ways and raise concern that virus spread and evolution could result in greater virulence (disease severity).

Enhancing viral transmission

The Gly^{614} (G614) mutation in spike (S) increases ordering of the 630 loop compared with wild-type Asp^{614} (D614). This prevents the premature S1 shedding often seen with wild-type S proteins, ensuring that more S protein remains in a fusion-ready "one-up" state, with one receptor binding domain (RBD) exposed within the trimer, ready to bind angiotensin-converting enzyme 2 (ACE2) on host cells, increasing infection efficiency.




3-RBD-down
RBD-up states
3D-1
340,327 particles
2.8 Å
3D-2
214,521 particles
3.1 Å
1U-1
412,536 particles
2.8 Å
1U-2
346,973 particles
2.8 Å
3D-3
276,950 particles
2.9 Å
3D-4
193,850 particles
3.2 Å
1U-3
367,939 particles
2.9 Å
2U
179,963 particles
3.0 Å

Extensive intrahost evolution of SARS-CoV-2 has been reported in at least five individuals with protracted infection because of immune impairment from therapy for hematologic malignancies or autoimmunity (*3*–*7*). They had active SARS-CoV-2 infection for an average of 115 days before clearing the infection or succumbing to COVID-19. Each patient also had at least one convalescent plasma (CP) treatment (intravenous transfusion of blood plasma from a donor who has recovered from COVID-19) and/or monoclonal antibody therapy. Some of these individuals were shedding high titers of SARS-CoV-2 at the time of discharge from hospital or before death, indicating the potential for transmission. SARS-CoV-2 variants from two of these patients had up to fivefold reduction in neutralization sensitivity to CP (*3*, *7*).

(*3*, *7*). Although these are case studies in immunocompromised individuals, they raise concern because the deletions of amino acids 69 to 70 (Δ69–70), Δ141–144, or Δ242–248 in S1 were observed in four out of the five infections (*3*, *5*–*7*); the N501T (Asn^{501}→Thr) or N501Y (Asn^{501}→Tyr) mutations were seen in two out of the five (*5*, *6*); and the E484K (Glu^{484}→Lys) and Q493K (Gln^{493}→Lys) mutations in the RBD of one infection also arose in antibody-resistant viruses after in vitro selection.

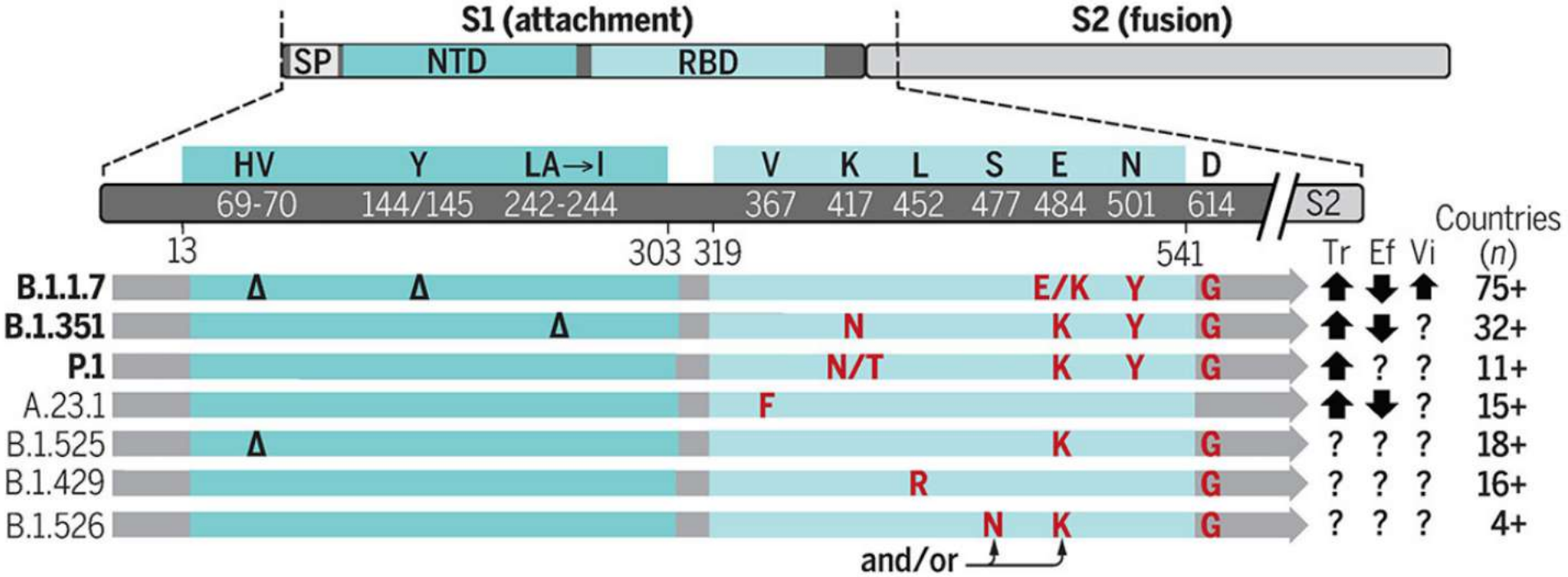
These reports preceded the detection of three major circulating variants—B.1.1.7, B.1.351, and P.1—which all contain at least eight single, nonsynonymous nucleotide changes, including E484K, N501Y, and/or K417N (Lys^{417}→Asn) in the ACE2 interface of the RBD. There are also various deletions in the amino (N)-terminal domain (NTD) of S1 in B.1.1.7 and B.1.351. Although most of the mutations in these variants were observed in a minor fraction of SARS-CoV-2 sequences during the first year of the pandemic, including K417N, E484K, and N501Y, there is no evidence to suggest that these variants were created through sequential addition of each substitution during interhost transmission.

The individual phenotypic effects of the mutations in S1 are incompletely understood, but some initial clues are emerging. Substitution at position Asn^{501} with Thr or Phe increases affinity for ACE2 binding (*9*), and Tyr^{501} increases infectivity and virulence in a mouse model (*10*). Some circulating variants may have reduced sensitivity to neutralizing antibodies that bind to the RBD directly (attributed to triple substitutions of key amino acids in the RBD at the ACE2-binding interface: Lys^{417}, Glu^{484}, and Asn^{501}) or to the NTD (conformational changes in the NTD are required for ACE2 attachment). More studies to correlate viral genotype and phenotype are needed.

It is possible that mutations that reduce neutralizing antibody binding, such as E484K, may require compensatory mutations that restore infectivity, such as N501Y. There appears to be convergent association of mutations such as the triple RBD mutation (Lys^{417}, Glu^{484}, and Asn^{501}) that evolved in two distinct lineages (B.1.351 and P.1).

Mutations and deletions in the spike protein

Currently, B.1.1.7, B.1.351, and P.1 are the major circulating variants of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2); others are emerging. The spike S1 subunit contains an amino (N)-terminal domain (NTD) and receptor-binding domain (RBD), which mediate host receptor recognition and contain epitopes for antibody binding. Deletions (NTD) and substitutions (RBD) in S1 can affect transmissibility (Tr), vaccine efficacy (Ef), and virulence (Vi). Additional mutations that define the variants can be tracked at (*8*). SP, signal peptide.

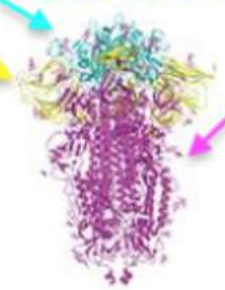


A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; N, Asn; R, Arg; S, Ser; T, Thr; V, Val; Y, Tyr.

Because only a few SARS-CoV-2 mutations were in circulation during most of 2020, it is likely that the three major variants are the result of selective pressures and adaptation of the virus during prolonged individual infections and subsequent transmission. All the case reports of individuals with extensive intrahost SARS-CoV-2 evolution indicated that they had been treated with suboptimal neutralizing antibodies (that is, the CP treatment did not neutralize the entire virus population). Whether or not antibody therapy played a role, it is likely that the same variants or variants containing new mutations will continue to emerge in different geographic locations as the result of intrahost selection and subsequent transmission. Indeed, other variants have been reported with multiple mutations in S1, including the lineages B.1.526 (detected in New York) and B.1.429 (which originated in California) containing a substitution in the RBD that is distinct from other variants; and B.1.525 and A.23.1 that are thought to have originated in Nigeria and Uganda, respectively

Spike chimera 1

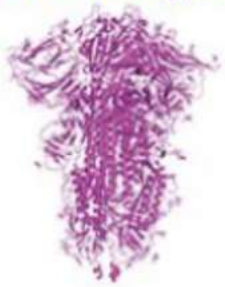
KU3-1 NTD/SARS-CoV RBD/SARS-CoV-2 S2



Spike chimera 2

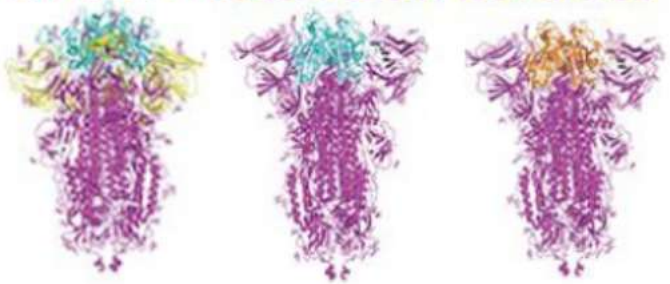
SARS-CoV-2 RBD/SARS-CoV S1 and S2

SARS-CoV-2 wild type furin knockout



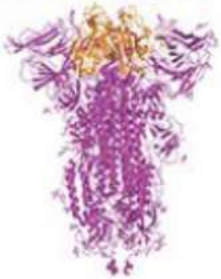
Spike chimera 3

SARS-CoV RBD/SARS-CoV-2 S1 and S2



Spike chimera 4

RsSHC014 RBD/SARS-CoV-2 S1 and S2



Norovirus capsid