Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Review

The spike glycoprotein of SARS-CoV-2: A review of how mutations of spike glycoproteins have driven the emergence of variants with high transmissibility and immune escape

Pedro F.N. Souza a,b,*, Felipe P. Mesquita b, Jackson L. Amaral a, Patrícia G.C. Landim a, Karollyny R.P. Lima a, Marília B. Costa b, Izabelle R. Farias b, Mônica O. Belém c, Yago O. Pinto d, Heline H.T. Moreira e, Ilana C.L. Magalhaes e, Débora S.C.M. Castelo-Branco f, Raquel C. Montenegro b, Claudia R. de Andrade c

a Department of Biochemistry and Molecular Biology, Federal University of Ceará, Brazil
b Drug research and Development Center, Department of Physiology and Pharmacology, Federal University of Ceará, Brazil
c Laboratory of Translational Research, Christus University Center, Fortaleza, Ceará 60192, Brazil
d Medical Education Institution Idomed, Canindé, Ceará

e State University of Ceará, Fortaleza, Ceará, Brazil
f Department of Pathology and Legal Medicine, Postgraduate Program in Medical Microbiology, Group of Applied Medical Microbiology, Federal University of Ceará, Fortaleza, Ceará, Brazil

ARTICLE INFO

Keywords:
Coronaviruses
Spike proteins
SARS-CoV-2
RBD
Mutations

ABSTRACT

Late in 2019, SARS-CoV-2 (severe acute respiratory syndrome coronavirus-2) emerged, causing an unknown type of pneumonia today called coronaviruses disease 2019 (COVID-19). COVID-19 is still an ongoing global outbreak that has claimed and threatened many lives worldwide. Along with the fastest vaccine developed in history to fight SARS-CoV-2 came a critical problem, SARS-CoV-2. These new variants are a result of the accumulation of mutations in the sequence and structure of spike (S) glycoprotein, which is by far the most critical protein for SARS-CoV-2 to recognize cells and escape the immune system, in addition to playing a role in SARS-CoV-2 infection, pathogenicity, transmission, and evolution. In this review, we discuss mutation of S protein and how these mutations have led to new variants that are usually more transmissible and can thus mitigate the immunity produced by vaccination. Here, analysis of S protein sequences and structures from variants point out the mutations among them, how they emerge, and the behavior of S protein from each variant. This review brings details in an understandable way about how the variants of SARS-CoV-2 are a result of mutations in S protein, making them more transmissible and even more aggressive than their relatives.

1. Introduction

SARS-CoV-2 belongs to the Coronavirus family, consisting of positive-sense, single-stranded RNA viruses that are lipid-enveloped [1]. In December 2019, SARS-CoV-2 infected its first patient, starting what would be a new pandemic established by the World Health Organization (WHO) [2,3]. The ongoing outbreak showed how devastating a virus can be even though SARS-CoV-2 is not that dangerous compared to other deadly viruses [4–8].

Unlike SARS-CoV-1 and MERS-CoV outbreaks, which were more endemic than a pandemic, SARS-CoV-2 quickly became global, based on high transmissibility [2,3,9,10]. At the beginning of the pandemic, COVID-19 was compared to other viral diseases caused by coronaviruses, even the flu (Table 1). Earlier in the outbreak, even some scientists did not give the necessary attention to COVID-19. However, as a completely new disease, COVID-19 showed that it cannot be neglected [11–15].

For example, in comparison with flu, COVID-19 has higher values of $R_0$, CRF, incubation time, and hospitalization rates (Table 1). Those results indicate that COVID-19 is not only just a strain of flu and deserves...
attention [16–19]. Compared to diseases caused by SARS-CoV and MERS-CoV, COVID-19 presents low CRF value, so that SARS-CoV and MERS-CoV have higher mortality than SARS-CoV-2 (Table 1). The infection caused by SARS-CoV-2 has a broader clinical spectrum, ranging from asymptomatic infection to severe viral pneumonia with respiratory failure and death [18,20,21].

The high $R_0$ of SARS-CoV-2 (Table 1) is probably due to the S protein's higher affinity (20 times more) than S protein from SARS-CoV-1 and MERS-CoV. The accumulation of favorable mutations in the S protein from SARS-CoV-2 made it more effective in recognizing human receptors than the spike from [14,22–26]. The mutations' accumulation in S proteins changed the S protein among different coronaviruses. It also led to genetically different lineages of SARS-CoV-2, called variants, that emerged and spread worldwide since the beginning of SARS-CoV-2 (Fig. 1) [27–31]. Today, there are two definitions of the variants of SARS-CoV-2: 1) variants of interest – VoI; and variants of concern – VoC (Table 2). VoIs possess mutations that improve SARS-CoV-2 fitness. VoCs arise from genetic changes that increase SARS-CoV-2 transmissibility and virulence, and can decrease the effectiveness of vaccines and treatments (Fig. 1) [27]. In the case of VoCs of SARS-CoV-2, these variants accumulated mutations on the S protein that either enhanced the transmissibility or the escape from the immune system. Sometimes these mutations are shared between the variants, while in other cases they are not (Fig. 1).

To the best of our knowledge, so far, no study has been published containing all the relevant information about SARS-CoV-2 VoCs, discussing how the mutations accumulate and affect the fitness of the S protein, and hence SARS-CoV-2. Based on that, this review is focused on the discussion and tracking of the mutations in the S protein that have led to the emergence of VoCs (Table 1). Additionally, we aim to understand the contribution of mutations to the S protein's three-dimensional structure and how these alterations affect the transmissibility and virulence of SARS-CoV-2 VoCs compared to the wild-type version that first came from Wuhan.

2. General aspects of SARS-CoV-2

2.1. Molecular biology of SARS-CoV-2

SARS-CoV-2 is a single-stranded, non-segmented, positive-sense RNA with both 5'cap and a tail at 3'end [32,33]. The SARS-CoV-2 RNA is similar to cellular mRNA, which is an essential evolutionary advantage, by mimicking the cellular mRNA, directly driving viral RNA for translation [34]. The positive-sense RNA of SARS-CoV-2 allows the infection in any permissive cell [33,35–37]. This characteristic of SARS-CoV-2 and other coronaviruses led Pfizer and BioNTech to produce the first mRNA vaccine (BNT162b2). BNT162b2 is a nucleoside-modified positive sense-RNA encoding the full-length membrane-anchored SARS-CoV-2 S protein enclosed in a lipid nanoparticle. After the shot, the skin promptly dissolves the positive-sense RNA and is translated by cell machinery of protein synthesis to viral S protein. Furthermore, S protein is exposed and externalized to be recognized by the immune system leading to the production of anti-SARS-CoV-2 antibodies [38].

During the SARS-CoV-2 replication cycle, the first protein produced is a large replication complex. Then RNA, which has a ribosomal frameshifting site [39], undergoes this frameshifting event. This event leads to the displacement of ribosomal frames, also known as translation.
recording, resulting in multiple protein production from a single mRNA molecule [39–44]. Moreover, the RNA serves for protein synthesis, RNA replication, and assembly, producing new virus particles [45]. The SARS-CoV-2 genome produces 16 non-structural proteins (nsp), including PL proteinase, 3CL protease, RNA-dependent RNA polymerase (RdRp), and helicase. The genome of SARS-CoV-2 also produces five structural envelope proteins (E), a membrane attached protein (M), nucleocapsid protein (N), and S protein (Fig. 2A).

### 2.2. Origin and evolution of SARS-CoV-2

SARS-CoV-2 is grouped in the B lineage of Beta-coronaviruses, the same as SARS-CoV [46–48]. Both SARS-CoV and SARS-CoV-2 share 79% of genomic identity, allowing similar structural, genetic, and pathogenic characteristics [48]. Despite many gaps in knowledge of SARS-CoV-2 origins, the WHO has started an investigation of SARS-CoV-2 [49–51].

This investigation is essential to avoid unbiased speculation that SARS-CoV-2 was developed at laboratory. Today, there are two front lines of investigations. First, the SARS-CoV-2 resulted from a laboratory incident at the Wuhan Institute of Virology (WIV). Second, whether SARS-CoV-2 came out from Wuhan markets, mainly the Huanan market, which is involved in trading of thousands of wild animals, including species with high risk, such as civets and raccoons [46]. The positive results for SARS-CoV-2 in environmental samples from the Huanan market support this theory [49–51].

Rather than speculation, here only findings based on scientific data will be discussed. Although there are many gaps in knowledge about the ribosomal frameshifting site’s influence on SARS-CoV-2’s origin, some studies have proposed a plausible explanation. Zhou et al. [52] suggested that SARS-CoV-2 very likely came from an ancestor of naturally SARS-like coronavirus from bats. The analysis of the genomic sequence of SARS-CoV-2 provided a clue to the evolutionary aspect of SARS-CoV-2.

### Table 2

Variants of SARS-CoV-2.

| Type               | Name     | Pango classification | Origin   | Type               | Name     | Pango classification | Origin   |
|--------------------|----------|----------------------|----------|--------------------|----------|----------------------|----------|
| Variants of interest | Lambda   | C.37                 | Peru     | Variants of concern | Alpha    | B.1.1.7               | United Kingdom |
|                    | Mu       | B.1.621               | Colombia |                    | Beta     | B.1.351               | South Africa   |
|                    | Epsilon  | B.1.427               | United States |                    | Gamma    | P1                    | Brazil                |
|                    | Kappa    | B.1.617.1             | India    |                    | Delta    | B.1.617.2             | India            |
|                    | Eta      | B.1.525               | United Kingdom |                    | Omicron  | B.1.1.529             | Discovered on South Africa |
|                    | Lota     | B.1.526               | United States |                    |          |                       |          |
|                    | Zeta     | P2                    | Brazil   |                    |          |                       |          |

Fig. 2. Schematic representation of the SARS-CoV-2 particle and spike protein. (A) SARS-CoV-2 particle closing the viral genome. (B) Spike protein consists of the S1 (holding the NDT and RBD regions) and S2 (holding the FP, HR1 and HR2 regions) units. (C) The two conformations assumed by Spike protein. The closed conformation with RBD inside the structure and the open conformation with RBD exposed to recognize ACE2 receptor. Created with www.BioRender.com.
Genetic analysis revealed an identity of 96% between SARS-CoV-2 and RatG13 (from *Rhinolophus affinis*). These data support the idea that bats are reservoirs of SARS-CoV-2 progenitor's and SARS-CoV-2 ancestors' spillover from bats to humans. It is intriguing that even with higher genetic similarities, the RaTG13 S protein diverges in the RBD region from the S protein from SARS-CoV-2, suggesting it may not bind to human ACE2 [53]. Moreover, based on RBD sequencing and host-receptor interactions, Pangolins (*Manis javanica*) were considered a possible reservoir based on high similarity with RBD of S protein from SARS-CoV-2 [14].

Unexpectedly, the analysis revealed that S protein coronaviruses from neither pangolins nor bats possess the polybasic cleavage site, essential for SARS-CoV-2 infectivity and host range [54]. Mutational analysis revealed the presence of insertions and deletions close to the S1-S2 junction of S protein from SARS-CoV-2. This change indicates acquisition of polybasic cleavage sites by SARS-CoV-2, established as a new feature in the SARS-CoV-2 genome, had improved transmission, species spillover, and human-human transmission [55–58]. This gain of function by S protein from SARS-CoV-2 may lead to higher affinity, which is 20 times that compared to S protein from SARS-CoV [59]. This improved S protein led to a higher spread of SARS-CoV-2, reaching pandemic dimensions four months after the first case [5,11,60,61]. To date, COVID-19 caused by SARS-CoV-2 has infected more than 245 million people and claimed more than 5 million lives.

### 2.3. Pathogenicity of SARS-CoV-2

The improvement of S protein in SARS-CoV-2 changes aspects of its infection. SARS-CoV-2 contacts are hosted by contaminated respiratory air droplets and human airway epithelium cells [62,63]. Zhu et al. [63] showed that although sharing the same receptor (ACE2) to reach cells, S protein from SARS-CoV-2 is more efficient in infecting human airway epithelium cells than human coronavirus NL63 (HCoV-NL63) and SARS-CoV. SARS-CoV-2 entries in cells interact, causing S protein binding with specific cell receptors, ACE2 and TMPRSS2 [64]. In summary, the RBD present in the S1 subunit from S protein binds to ACE2 in the host cell. Then, the TMPRSS2 catalyzes the cleavage of S protein on the S1-S2 domain to start the membrane fusion process and viral entry [65]. Later, this mechanism will be detailed in the section about S protein.

SARS-CoV-2 infection can result in three situations: 1) asymptomatic, 2) mild symptoms, and 3) severe acute pneumonia and death [66–69]. In situations 2 and 3, the symptoms vary from mild, such as fever, chills, fatigue, cough, dyspnea, chest pain, myalgia/arthritis, diarrhea, nausea and vomiting to severe such as infections in the lungs, intestinal tract, pharynx, heart, kidney, liver, brain, blood, inflammatory reactions and alveolar damage [66–69]. Most of the problems during COVID-19 are caused by an overreaction of the immune system. A process called "cytokine storm" produces a high level of proinflammatory cytokines such as growth factor-β1 (TGF-β1), tumor necrosis factor-α (TNF-α), interleukin-1β (IL-1β), IL-6, IL-2, IL-7, IL-10, and granulocyte colony-stimulating factor (G-CSF), leading to local inflammation and multiple organ functional failure [68,70–73].

## 3. The S protein

Before moving forward with the discussion about S protein in VoCs I, it is essential to understand how this protein works. Like other viral attachment proteins, S protein is synthesized and deposited in viral particles in inactive form. Thus, during the viral infection, it is activated by a host protease, which releases the S1 and S2 subunits, activating them and allowing the membrane fusion and virus entry as an inactive precursor present in the viral envelope. The S protein is cleaved during cell infection by a cellular protease, releasing [64,74,75]. An exciting characteristic of S protein is the high level of glycosylation. Before post-translational processing, S protein is present in amounts ranging from 128 to 160 kDa. After carbohydrate insertion, mostly by N-links, the size of S protein increases to 180–200 [76–78]. It has been hypothesized that the biological function of glycosylation in the S proteins is to evade the host immune system during infection [76–78]. S protein has an N-terminal region that is about 90% of protein length exposed to the outside of the viral particle, a transmembrane section, and an inside C-terminal region [79]. Even at the monomeric stage with 1273 amino acid residues, S protein reveals its size [80], being divided into S1 (1-685) and S2 (886-1273).

The S1 subunit is involved in host cell recognition and attachment to it. The receptor-binding domain (RBD) responsible for binding to the ACE2 is hosted by the S1 subunit [58,75,81]. By holding the RBD domain, the S1 subunit is the most likely target for the action of neutralizing antibodies and developing drugs [58,64,65,74,81]. Additionally, S1 contains the N-terminal (NTD) and C-terminal (CTD) domains, both involved in the RBD recognition. For comparison, the SARS-CoV-2 S1 CTD has an insertion of 21 amino acid residues (21 as) directly involved with the ACE2 interaction compared to SARS-CoV S1 CDT. The amino acid residues F486 and E484 from SARS-CoV-2 S1 CTD determine the strength of the interaction with ACE2 receptor compared to I472 and P470 from SARS-CoV S1 CTD [58,81].

The S2 subunit is involved in membrane fusion and virus entry in cells. To develop this function, the S2 subunit hosts three critical domains: fusion peptide (FP, 798–806), heptad repeat sequence 1 (HR1) (912–984 residues), and heptad repeat sequence 2 (HR2) (1163–1213). FP is a small apolar fragment composed of Gly and Ala, allowing the interaction with the membrane, densely conserved in the coronavirus family [21,82,83]. Two FP domains are found in two repeated sequences, HR1 and HR2. These amphipathic sequences with hydrophobic and hydrophilic amino acid residues allow all types of interactions with phospholipids in the membrane. HR1 and HR2 interact to form a domain fusion core region [84] that drives the S2 subunit to the cell membrane and initiates the fusion process [82,84–86].

Many other viruses present S-like proteins critical to receptor recognition, cell attachment, membrane fusion, and virus entry. HIV and Ebola possess S-like proteins with other names [64,69,81,87–95]. In SARS-CoV-2, the S protein in its trimeric form is found exposed outside the viral particle with an extension of 15 nm playing an essential role in viral infection. Exposed in the viral particle and before interacting with ACE2 receptor, the S protein is found in inactive form. The S protein possesses two conformational moments (Fig. 2C), the open and closed states.

This conformational flexibility is based on the presence of NTD and RBD. The RBD is enclosed inside the S protein structure in the closed state and cannot interact with the ACE2 receptor (Fig. 2C). The closed conformation is also known as the perfusion stage [93,96,97]. At this stage, the RBD is found in a pocket inside the S protein structure, covered by the NTD region. At this moment, the S1 stays on the top of the S2 subunit to prevent early conformational change to the post-fusion structure [98,99].

Some elegant studies have shown that a proteolytic cleavage of S protein changes from perfusion to post-fusion stage. This cleavage of S protein is made by a cellular protease attached to the cell membrane. The protease is TMPRSS2 and happens at a polybasic cleavage site on S protein. Interestingly, S protein from SARS-CoV-2 has a higher cleavage than S protein from SARS-CoV. Higher levels of S protein cleavage might be involved in the faster entrance in cells and thus faster replication of the virus [64,74].

The RBD is pulled out of the S protein structure during the open state and can recognize and interact with the ACE2 receptor (Fig. 2C) [64,65,100]. After the RBD and ACE2 receptor interaction, the complex RBD-ACE2 initializes the conformational changes of the S protein structure to the fusion state (post-fusion). The S2 subunit and membrane control this reaction. Membrane fusion leads to SARS-CoV-2 entry in cells [64,65,86,97–100]. Right after SARS-CoV-2 reaches the cytoplasm, the infection process starts [80,81,101,102].
4. How mutants of S protein led to VoCs

4.1. Mutation on coronavirus genome

As described above, viruses classified in the coronavirus family have the giant single-stranded RNA genomes. The complexity of the genome requires a complex mechanism to repair possible mistakes in the replication process. So, coronaviruses possess a massive polymerase complex, and a multifunctional protein (nsp14) proofreading the Exo N domain. In contrast, this proofreading function is not found in other RNA viruses with smaller genome sizes [93]. The domain with residues compatible with Exo N activity has been described in other proteins. Genetic diversity in SARS-CoV-2 is very likely low [103–105]. However, natural selection can cause rare but favorable mutations. For example, mutations in the influenza virus during each year's flu season leads to a complex balance between immunological resistance developed across populations and mutations, providing antibody resistance [61], which requires the development of new influenza vaccines every season [106]. In a recent study, Sehra et al. [107] suggested the establishment of SARS-CoV-2 as a seasonal infection like flu. However, the long duration of the outbreak can lead to the accumulation of immunologically relevant mutations of the SARS-CoV-2 genome even in the vaccinated population.

4.2. Could VoCs overshadow the breakthrough of vaccines?

To have an overview of the impact of mutations that generate VoCs on vaccine efficacy as revealed by experiments with SARS-CoV-2 itself or pseudoviruses expressing combined or individual mutations on S protein that led to VoCs, sera of 20 post-vaccinated volunteers immunized with BNT162b2 (Pfizer–BioNTech) or mRNA-1273 (Moderna), both mRNA vaccines, presented levels of anti-spike IgM and IgG and RBD-specific antibodies similar to people that were naturally infected by SARS-CoV-2 [108–111]. The results revealed a reduction in the neutralization of SARS-CoV-2 or pseudoviruses carrying mutations compared to the Wuhan isolate.

In another experiment, the effect was evaluated of a single N501Y mutation on neutralization. Results showed that sera from 20 participants in a trial of BNT162b2 presented a modest reduction in the neutralization of virus, with N501Y mutation [109,110,112,113]. Regarding the virus holding the E484K mutation, results revealed a 6.7-fold reduction of the neutralization of post-vaccination sera from people vaccinated with the BNT162b2 [114]. Experiments with Alpha and Beta variants showed a higher neutralization action against live-viral antibodies from sera of people vaccinated with ChAdOx1 nCoV-19 (Oxford–AstraZeneca) [115]. Gamma variant was able to escape from antibodies produced by BNT162b2 (5.8-fold) and mRNA-1273 (4.5-fold and 2.9-fold, respectively), reducing the neutralization of vaccines, respectively, by 6 and 4.5-fold post-vaccination [116]. Both Beta and Gamma VoCs share the E484K; suggesting this mutation could be a determinant for low neutralization of post-vaccinated sera.

The NVX-CoV2373 (Novavax) S protein-based vaccine showed 95.6% efficacy against the wild type. However, this efficacy dropped to 85.6% and 60.0%, respectively, for the Alpha and Beta variants [117]. Likewise, the vaccine JNJ-78436735 (Johnson & Johnson/Janssen) showed efficacy of 72% for wild types, but this number dropped to 57% for the Beta variant [118]. These data indicate the Beta variant has accumulated mutations that reduced the efficacy of NVX-CoV2373 and JNJ-78436735, but these vaccines are still clinically efficacious.

Fig. 3. Scheme presenting the mutations on variants and structural alignments of the variants compared to wildtype version Spike protein from SARS-CoV-2. (A) Comparison of Spike protein sequences from wildtype version of SARS-CoV-2 and mutations on Spike protein presenting in the variants D614, Alpha and Beta. (B) structural analysis of Spike protein from wildtype version of SARS-CoV-2, (C) Spike protein from Alpha variant and (D) Spike protein from Beta variant. RMSD analysis indicates changes in the structure.
5. S protein holding D614G mutation

The isolate of SARS-CoV-2 hosting the SD614G (Fig. 3A) emerged and spread globally in January to March 2020 [119,120]. Besides showing up early in 2020, the D614G is a well-conserved substitution presented by all new SARS-CoV-2 VoCs that came after suggesting it as a well-established successful mutation [119]. SD614G brought many benefits to SARS-CoV-2 fitness, which will be discussed, such as enhanced replication and infectivity. It facilitates S protein incorporation into the virion, increases virion stability, leads to higher viral loads in COVID-19 patients, modulates glycosylation at the nearby N616 site, but has higher infectivity and transmissibility of SARS-CoV-2 [119–123].

At the molecular level, the D614G mutation increases the stability and density of the S protein. It decreases the S1 subunit shedding, keeping both S1 and S2 subunits more closely supported by stronger intermolecular association, leading to a highly stabilized S1–S2 interface in SD614G. It was known that S1 shedding limits the infectivity of the original SARS-CoV-2 (without SD614G) by releasing S1 from S2 during viral fusion. The D614G mutation fixed this problem by strengthening the intramolecular interaction between S1 and S2, blocking S1 dissociation and thus virus infectivity and transmission [119–123].

6. Alpha variant of SARS-CoV-2

In December 2020, a variant that spread quickly and increased the risk of death was detected in the United Kingdom (UK). The variant derived from the SARS-CoV-2 20B clade was confirmed as B.1.1.7, now known as the Alpha variant. In the early investigation stage, increased transmissibility of up to 71% over the previously circulating strains of SARS-CoV-2 [124,125] was estimated. The high transmissibility of the Alpha variant has been described in several studies in different countries, with a transmissibility range between 50 and 80% [126–128]. In vitro and in vivo experiments supported the observed behavior of the Alpha variant. It was demonstrated that the Alpha variant can multiply and shed more effectively in the nasal cavity of hamsters than other variants, even at low viral loads and short time of exposure [129]. Based on that, the Alpha variant was defined as VoC in December 2020. It was later defined as a variant, being monitored (VMB) because it had no longer been detected or was circulating at very low levels [130].

In January 2021, 45 countries reported the Alpha variant's presence. A group of British scientists proposed that the Alpha variant is responsible for the third wave of the pandemic in Scotland, associated with severe cases of COVID-19. The study results showed that the Alpha variant of SARS-CoV-2 has 5 times higher permeability than other co-circulating variants, hence its rapid spread in Scotland [131].

In a UK case-control study of 54,906 participants who tested positive for SARS-CoV-2, in comparison with the Alpha variant and previously circulating variants reported a mortality hazard ratio associated with infection with VOC-202012/1 compared with infection with previously circulating variants (hazard ratio of 1.64 [95% CI 1.32–2.04]) [132]. The Biwako Ohashi Hospital in Japan described a specific sign called red face after Alpha variant infection, which may be predictive of a sudden deterioration of patients [133].

Interestingly, in September 2021, a French group had already identified the SARS-CoV-2 Alpha variant in cats and dogs in France [134]. Recently, scientists in the UK identified cases of severe myocarditis in cats and dogs associated with the Alpha variant that crossed the interspecies line in that country, raising questions regarding its potential pathogenicity in these animals [135]. The Alpha variant results from the accumulation of 17 mutations compared to the original SARS-CoV-2 virus discovered in Wuhan, China. Here, as before reported, we focus on the mutations accumulated in the S protein (Fig. 3A). These mutations are expected to lead to expanded spread and increased severity of the associated disease, requiring more effective vaccines, therapeutic drugs, diagnostic tools, and other public health measures.

The Alpha variant has eight spike mutations, including two deletions in the S protein (Fig. 3A). These mutations led to two deletions, both in the NDT region: 1) 69–70del (i.e., a deletion of 6 bases coding for histidine and valine at positions 69 and 70, respectively, in the viral S gene); and 2) 144del (deletion of Tyr residue). Usually, deletions of SARS-CoV-2 result in escape from the immune system. However, the 69–70del has a different purpose. The 69–70del is very likely a genetic compensation. As we will discuss further, mutations of RBD to escape from the host's immune system led to reduced affinity for ACE2 receptors. So, 69–70del can increase the affinity of S protein for ACE2 receptors [136]. Kemp et al. [136] revealed in vitro that 69–70del is not involved in the escape of antibodies from convalescent sera but does increase infectivity, incorporation of cleaved S protein on viral membrane, and spike infectivity. The authors discussed the idea that this deletion came about to compensate for small infectivity defects induced by RBD mutations N501Y, N439K, and Y453F in the S protein sequence [136]. The 144del is a deletion of Tyr residue. It is supposed that deletion changes the NDT loop based on the high volume of Tyr residue. This alteration in the NDT loop might indicate that neutralization by 144del is more efficient in reducing neutralization by NTD-specific neutralizing antibodies (9 of 10; 90%) than RBD-specific neutralizing antibodies (5 of 31, 16%) [110,137–140].

The next mutation in the S protein from Alpha is N501Y (i.e., an asparagine to tyrosine amino acid substitution at position 501 in the viral S gene) (Fig. 3A). It is the first time this mutation was detected in SARS-CoV-2. The 501 position is at the RBD–ACE2 interface. N501Y forms part of the binding loop in the contact region of hACE2, forming a hydrogen bond with Y41 in hACE2. It also stabilizes K353, one of the virus-binding hotspot residues in hACE2. The N501Y substitution is an advantage to variants holding it, maybe because it is present in Alpha, Beta, and Gama variants (Figs. 3 and 4). The substitution of an asparagine (polar) with a tyrosine (hydrophobic) increases the affinity of RBD for ACE2a [141,142]. Molecular investigations showed that the N501Y mutation of the SARS-CoV-2 variant identified in the UK and South Africa evolved separately [143]. N501Y mutation slowed the dissociation of the RBD from the ACE2 receptor. Additionally, it seems that N501Y reduces neutralization by antibodies [110,124].

The effect of a mutation that replaces an alanine with an aspartate residue at position 570 in the S protein remains unclear. It has only been detected in the Alpha variant (Fig. 3A). The same happens with the T716I, S982A, and D1118H mutations; they have only been found in the Alpha variant, and their functions are still unclear (Fig. 3A). However, the mutation replacing proline with arginine (P681H) increases the positive net effect of S protein right before the 685 position, which is the furin cleavage site (Fig. 3A). It has been proposed that P681H enhances the cleavage by a furin protease, which is involved in the activation of membrane fusion during viral infection [64,144,145].

All those mutations change the S protein, and root mean square deviation (RMSD) calculation revealed the atomic position on the S protein of variants compared with the wild-type S protein from the Wuhan isolate (Figs. 3 and 4). The RMSD calculation provides a clue about the alterations in protein structure through a score: a score of 0 indicates the atoms from both structures are in the same position, which means no alteration in the 3D structure. Any score different from 0 indicates the atoms are in different positions, and the higher this value is, the greater the differences are between structures. The S protein from the Alpha variant has accumulated mutations on the HR1 and HR2 portions of the S2 subunit, leading to conformational changes in those regions, which are essential to membrane fusion.

Research has shown that vaccines still provide reliable protection against the variant, although they may be less effective. Different studies have described the impact of the Alpha variant on vaccine efficiency. The BNT162b2 vaccine (Pfizer) showed 0.81- to 1.46-fold reductions in geometric mean titers (GMTs) against mutations in B.1.1.7 and B.1.351 variants relative to the parental virus from Wuhan [112,113]. The mRNA-1273 vaccine (Moderna) demonstrated no reduction against the
B.1.1.7 variant [111]. ChAdOx1-nCoV-19-elicited sera with 9-fold lower \textit{in vitro} neutralization activity against the B.1.1.7 variant [114]. Based on these findings, scientists suggest that in countries with low vaccination coverage, where the Alpha variant is not yet dominant, there is an urgent need to improve health facilities to meet medical requirements of possible Alpha outbreaks.

7. Beta variants of SARS-CoV-2

The Beta variant (B.1.351) emerged in Nelson Mandela Bay, Eastern Cape, in October and November 2020 and soon became predominant in South Africa, overcoming the previously dominant variants B.1.1.54, B.1.1.56, and C.1 lineages [146]. By December of 2020 and January of 2021, the Beta variant had arrived in Botswana, France, Scotland, South Korea, Sweden, Switzerland, the United Kingdom, Austria, Belgium, Denmark, Finland, Germany, Ireland, Netherlands, Norway, Australia, Brazil, Canada, China, Japan, Taiwan, United States, and Zambia [147]. This fast spread led to over 95% of infections during and after the second epidemic wave, between November 2020 and February 2021 [142,146,148].

The Beta variant was initially classified as a new monophyletic cluster (501Y.V2), replacing the three main South African lineages (B.1.1.54, B.1.1.56, and C.1) that were circulating during the first epidemic wave [148]. The high transmissibility and increased infection duration explain why the Beta variant became predominant [149,150]. According to the mutational profile observed at the first sampling point, this lineage had accumulated, in addition to D614G, five other non-synonymous mutations in the spike protein (D80A, D215G, E484K, N501Y, and A701V). Three further spike mutations emerged by the end of November (L18F, R246I, and K417N). Additionally, the Beta variant has two deletions in the NDT region (Fig. 3A) [148].

The substitutions are: L18F, involving the exchange of a leucine (apolar amino acid) by a phenylalanine (a hydrophobic amino acid); D80A, the replacement of aspartic acid (negative) with alanine (apolar); D215G, replacement of an aspartic acid (negative) with a glycine (apolar); and R246I, involving replacement of an arginine (positive) with an isoleucine (apolar). Also, deletions in the NDT region (Fig. 3A) are involved in the escape from neutralization by antibodies such as NTD-binding mAbs30 and 4A8 [140].

Three of the spike mutations are in critical residues in the receptor-binding domain (RBD) (K417K, E484K, and N501Y) (Fig. 3A). The K417 position of the S protein has been designated an epitope of RBD of class 1 and 2 antibodies [109]. Moreover, very likely the alteration of this spot affects the ability of class 1 and 2 antibodies to bind and neutralize the RBD, preventing the interaction with the ACE2 receptor [141,142]. In the Beta variant, a substitution of lysine (positive) by asparagine (polar uncharged) is involved in the ability of the S protein from the Beta variant to escape from immune antibodies. Besides that, this substitution has a negative effect. It was noticed that K417N slightly reduces the affinity of RBD for ACE2, but it seems this collateral effect did not change the Beta variant’s fitness [142].

The substitutions E484K and on RBD possibly enhanced the binding affinity of the Beta variant for the ACE2 receptor. The E484K mutation is uncommon and interacts with the K31 hotspot residue of hACE2 [142,151]. The position E484 of the S protein indicates it is an immunodominant residue, which means the immune response is focused against it. So, mutation in this position could completely disarm the immune response toward a variant holding it. The Beta variant was the first VoC to show substitution at position 484, in which a proline (apolar) was replaced by lysine (positive). The E484K is involved in...
escape from antibodies in convalescent plasma (C121 and C144 monoclonal antibodies) [151-155]. Baum et al. [155] showed that a pseudovirus with the mutation E484K could escape neutralization by combining two monoclonal antibodies, REGN10989 and REGN10934, used to treat severe cases of COVID-19.

The mutation N501Y (see the Alpha variant section) is present in Alpha and Beta. It is involved in the binding phase of the S protein with ACE2, allowing an increase in affinity for ACE2 and modifying the structural mechanics of the RBD-ACE2 complex. The N501Y enables escape from the immune system but reduces the affinity of the S protein for the ACE2 receptor (see the Alpha variant section). However, in the Beta variant, the presence of K417N and E484K mutations slightly reduce the affinity for RBD, as they abolish the interfacial salt bridges that help bind RBD-ACE2 and stabilize the complex. Therefore, these two additional mutations in the Beta variant could overshadow in part the increase in the RBD-ACE2 caused by N501Y. However, those mutations facilitate the escape of the Beta variant from the neutralizing effect of antibodies. Based on that, it is feasible to suggest that the positive effect of K417N and E484K mutations is greater than the negative effect, enhancing the Beta variant’s fitness and thus being kept by the virus and even developed by other variants [50,156].

As discussed above, the Beta variant presents many non-synonymous amino acid substitutions. These amino acid changes alter the three-dimensional structure of the S protein. For example, N501Y replaces an asparagine with a tyrosine. The side-chain volume of asparagine is 114.1 Å³, compared to the side-chain volume of 193.6 Å³ of tyrosine. This difference in the size of lateral chains affects the S protein’s structure. To verify that, a root mean square deviation (RMSD) calculation of the atomic position on the S protein from the variants was performed in comparison with the wild-type S protein from the Wuhan isolate (Figs. 3). To perform the structural alignments, the 3D structures of S proteins of the Wuhan isolate (PDB ID: 6Z97) were deposited in the Protein Data Bank (PDB, https://www.rcsb.org/) and the database of the Beta variant (modeled in this study). The RMSD analysis revealed significant structural changes between Wuhan S protein and Beta variant S protein (Fig. 3D). It is possible to analyze in many positions the differences of the two proteins. The RMSD value for this alignment was 1.604 Å, indicating that both atomic positions are quite different. The greatest alteration was in regions such as NTD, the arm of S1, and RBD (Fig. 3D). These alterations are responsible for the differential interaction of the S protein with the ACE2 receptor, and thus the spread, infectivity, and severity of the disease.

Given the ability to escape from the immune system, the Beta variant overtook the Alpha variant in regions with naturally acquired immunity, by about 20% to 40%. To date, modest clinical trials have shown that two doses of the Pfizer-BioNTech vaccine had 75% effectiveness against Beta variant infection. In contrast, Novavax’s vaccine showed 89% efficacy in the UK compared to South Africa, with 60%. Similarly, the Johnson & Johnson vaccine trials reported lower levels of protection against moderate to severe COVID-19 in South Africa than in the United States. Meanwhile, South Africa ceased plans to roll out the Astrazeneca vaccine because clinical trials did not show protection against mild or moderate illness caused by the Beta variant [145,156,157]. Experiments using viruses expressing RBD with K417N + E484K + N501Y revealed a reduction of 2 to 3-fold in neutralization by antibodies in plasma from vaccinated individuals [111].

8. Gamma variant of SARS-CoV-2

Since SARS-CoV-2 became pandemic, it has evolved genetically leading to different variants (Fig. 1), with different genetic profiles, and thus different degrees of severity. The genomic sequencing of viral samples has been fundamental to detect the new variants of SARS-CoV-2 [25,65]. The new VoC, the P.1 lineage, known today as the Gamma variant, was detected in about 42% of the analyzed genomic sequences in samples from Manaus in December 2020 and presented important mutations with biological significance.

The Gamma variant was also identified in January 2021 in Brazilian travelers at Japanese airports [158]. After that, this variant was identified in different Brazilian regions [159-161]. Manaus, the capital of Amazonas State, was the most affected and became an epicenter of the outbreak of the Gamma variant, which quickly spread through the Solimoes River course, causing the collapse of public and private health systems [162]. Notably, the cases of Gamma variant in January represented 85.4% (41/48) of COVID-19 cases in Manaus [163]. In February 2021, it had spread in South America and was responsible for nearly 40% of COVID-19 cases [164-166]. These results revealed the high infectivity rate of this variant, which is estimated to be 2.6 times more infectious.

In February 2021, two cases of a new SARS-CoV-2 variant were described in Alagoas State [167]. One person came fromAmazonas, and the other was characterized as community transmission in Alagoas, confirming the presence of the Gamma variant [161]. According to data gathered by the FIOCRUZ-COVID-19 Genomics Surveillance Network of the Brazilian Ministry of Health, the Gamma variant was mainly responsible for Covid-19 cases in Brazil from February to July 2021. For comparison, in December 2020, 26% of deposited genomes were the Gamma variant (P.1); six months later, in June, 64.6% were the Gamma variant, and in August, 99% of the sequenced samples were the Gamma variant [168].

After the first detection on 23 December 2020 in Manaus [163], the Gamma variant was found in 71 countries. It was detected in South America, Central America, North America, Europe, Angola, Turkey, Japan, the Philippines and Australia. The Gamma variant was not detected following a logical geographical sequence concerning the first case, since it was detected in Uruguay, Canada, India and Australia in the same period. Brazil and the United States had the greatest number of Gamma infections globally [169]. The Gamma variant was classified as VoC based on its greater transmissibility, pathogenicity, and immune escape. The Gamma variant carries multiple mutations in the NTD and RBD domains [170,171].

The analysis of sequences revealed that the SARS-CoV-2 Gamma variant has 37 mutations, including 22 missense, 10 synonymous, three intergenic, one frameshift, and one in-frame shift mutation. In the spike protein, 12 missense mutations (L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I, and V1176F) were observed [163,171,172]. Three mutations (K417T, E484K, and N501Y) were also identified in the RBD region of the spike protein [158,170].

Although the emergence was in different in most countries, the Gamma and Beta variants share some important mutations in RBD, such as L18F, K417T, and E484K, in addition to D614G (Figs. 3 and 4 panel A). The N501Y is shared by the Alpha, Beta, and Gamma variants. The appearance of those mutations in different variants that emerged in different countries has two possible explanations: first, the selective pressure on SARS-CoV-2, probably induced by either natural or artificial immunity, drove the emergence of the mutation in different individuals; or second, it is a clear case of convergent evolution. The implication of these mutations (L18F, K417T, E484K and N501Y) has already been discussed in the Alpha and Beta sections.

In addition to shared mutations, some mutations are exclusive to the Gamma variant (Fig. 4A). Those mutations are T20N, P26S, D138Y, R190S, H655Y, close to the furin site, and T1027I in the junction between the HR1 and HR2 regions. The mutations T20N (replacing asparagine with a tyrosine) and P26S (replacing a proline with a serine) are located in the NTD supertype region. The NTD supertype is highly targeted by antibodies [137,140,163]. Studies by McCul1um et al. [137] and Suryadevara et al. [173] revealed that antibodies with higher neutralizing power preferentially attack the NTD supertype region. Thus, a mutation in this region helps the Gamma variant to escape from the immune system. Additionally, the mutation T20N provides an N site for glycosylation, which could create a shield, protecting the region from antibody recognition [174]. Although D138Y, R190S, H655Y, and
T10271 are non-synonymous mutations, it seems that their mutations have not affected or conferred any additional fitness to the S protein, or at best, their effects are still unclear. What is known is that those mutations are lineage-defining.

RMSD analysis comparing the S protein from the Gamma variant with the S protein from the Wuhan isolate produced a score of 1.060 Å, indicating the differences between S proteins (Fig. 4B and C). Compared to other variants, the S protein from the Gamma variant is the one closest to the S protein from the Wuhan isolate. The most exciting alterations in the 3D structure are in the NDT region (Fig. 4C), and the mutation T20N, which potentially increases the glycosylation, covering the entire region, preventing the antibodies’ attachment. The RMSD analysis suggests that the S protein from the Gamma variant has adapted its structure mainly to escape from the immune system. The Gamma variant is responsible for the first case of reinfection in Manaus, Brazil [174]. This observation indicates the Gamma variant has excellent ability to mitigate or escape from the immune system, which might be related to a higher mortality rate of the Gamma variant, estimated to be 1.7 to 2.4-fold higher than the rate of the Wuhan isolate [174].

The Gamma variant has a high attack rate even in people that have completed the vaccination regimen. By July 2021, the Gamma variant was predominant in French Guiana, threatening the hospital capacity by causing the third wave [175]. The reduced vaccine efficiency toward the Gamma variant was a surprise, because in vitro trials showed strong neutralization of antibodies in people vaccinated with the BNT162b2 vaccine [38,110–113,116]. The unexpectedly high attack rate of the Gamma variant could result from problems with the vaccine storage, but this is just speculation. A Gamma outbreak among persons fully vaccinated with the BNT162b2 vaccine shows the ability of the Gamma variant to escape from the immune system. This uncommon infection of the Gamma variant in vaccinated people demands surveillance and studies of BNT162b2’s effectiveness in regard to the Gamma variant.

9. Delta variant

First reported in India, in late December 2020, the double-mutated Delta VoC, a coronavirus variant formerly known as Indian VoC or B.1.617.2, is 40% more transmissible than the Alpha variant and two times more transmissible than the original Wuhan strain of SARS-CoV-2, leading to new lockdowns [176,177]. The Delta variant required only a few months to spread to 98 countries, becoming the dominant variant in all of them. Delta is the causative agent of more than 83% of COVID-19 cases in the U.S. and 91% of infections in the UK. The UK and many countries in Europe have faced increased new infections and deaths by the Delta variant, which rapidly became the dominant variant worldwide [178].

The advantage of this Delta variant is the combination of high transmissibility and immune system evasion, besides being less sensitive to neutralizing antibodies. Mutations of the S protein may have enhanced its ability to bind to the ACE2 receptor [179]. As discussed above, the S protein must be cut twice by host proteins to enter human cells. However, the Delta variant has earlier processing of the S protein, and newly formed viral particles emerge rapidly from infected cells, infecting other cells more efficiently [180].

The Delta variant has accumulated multiple mutations in the S1 subunit, specifically in the NDT and RBG regions. Mutations in NDT are related to escape from the immune system, and mutation in RBG seems to improve its stability and ability to bind to ACE2 and evade the immune system [181]. However, despite this knowledge, many unknowns surround the Delta variant. For instance, it is not yet understood for sure how Delta’s mutations lead to a supercharged variant [181].

According to the Center for Disease Control (CDC), Delta has exclusive signature mutations including T19R, V70F*, T95I, G142D, delE156-, R158G, A222V*, W258L*, L452R, T478K, P681R, and D950N (Fig. 4A) [181]. The Delta variant also has the mutations D614G, shared with all variants, and K417N, shared with the Beta and Gamma variants. The implication of D614G and K417N has already been discussed in this study. Delta variants presenting the K417N mutation shared with Gamma and Beta variants belong to lineages AY.1, AY.2, and AY.3, being called Delta plus variants (Figs. 3A and 4A) [177–180,182–185].

The mutations T19R, V70F*, T95I, G142D, delE156-, delF157-, R158G, A222V* and W258L* are all located in the NDT (Fig. 4) region, covering the NDT supersite, which is the target of anti-NDT neutralizing antibodies [109,116,137,140]. The NDT region has a stabilizing effect on the S protein [186] because it is targeted by potent neutralizing antibodies [137,173,186–188]. The NDT supersite is divided into five loops, named N1 to N5, covering different residue gaps in N1 (residues 14–26), N2 (residues 67–79), N3 (residues 141–156), N4 (residues 177–186), and N5 (residues 246–260) [186,187,189].

Delta has accumulated many non-synonymous mutations in the NDT supersite, changing its structure to reduce the neutralizing effect of anti-NDT antibodies. A meticulous analysis of the Delta variant’s mutations in the NDT region shows that T19R is in the N1 region, V70F* is in the N2 region, G142D and delE156- are in N3, and A222V* and W258L* are in the N5 region. The Delta variant showed three mutations in the RBG region (K417N, L452R, T478K). One of them, K417N, shared with the Beta and Gamma variants, has already been discussed. Here we focus on the two exclusive regions of the Delta variants L452R and T478K (Fig. 4A).

It was shown that the Alpha, Beta, and Gamma variants share the mutation N501Y, which is involved in escaping from the immune system. However, this mutation slightly reduces the affinity of RBG for the ACE receptor. It seems that the Delta variant has accumulated mutations to overcome that obstacle. The Delta variant lacks the mutation N501Y. In contrast, it has the mutations L452R and T478K, which are involved in the escape from anti-RBG antibodies and do not reduce the affinity of RBG for ACE2 [28,46,93,136,137,186,187,190,191]. The mutation L452R (changing a leucine for an arginine) is closely related to the escape of monoclonal antibodies [142]. For example, the monoclonal antibody Bamlanivimab lost its antiviral effect against the Delta variant, contrasting with its continuing activity against other variants [142]. The unique change of T478K in the RBG region of the Delta S protein was found precisely in the epitope region recognized by monoclonal antibodies known as Class 1 [192].

The mutation P681R, adjacent to the furin cleavage site (685), seems to be one of the greatest advantages of the Delta variant. The Alpha variant has a mutation at the same P681H site. This mutation increases the Alpha transmissibility by 40% in comparison with the original SARS-CoV-2 from Wuhan. The mutation replaces the proline residue with an arginine residue in the Delta variant at position 681. Both mutations in the Alpha and Delta variants lead to a less acidic environment, favoring the basic cleavage at the furin site. The fact is that arginine is the most basic ever, with a pK of the side chain of 12.5. Based on that, arginine provides a perfect environment for furin cleavage and activation of the S protein.

Furin cuts indicate the existence of more S proteins primed to invade human cells. For example, in SARS-CoV virions, only 10% of S protein is primed and can infect cells. In the Alpha variant, with P681H, the primed S protein is about 50%. In contrast, in the Delta variant, holding the mutation P681R, 75% of S proteins are primed and can infect human cells. The P681R is not just one more mutation; it is behind the fast spread of the Delta variant and probably is the essential advantage gained by Delta [180].

Liu and researchers reported that the S protein is cut more efficiently in the Delta variant than others [193]. The P681R at a furin cleavage site speeds up the infection of uninfected cells through plasma membrane fusion, three times more than the other variants [185].
The D950N (Fig. 4A) region is located at the interface of the S protein trimer. It is suggested that this mutation regulates S protein dynamics between monomeric and trimeric phases [194]. The RMSD analysis (Fig. 4D) revealed that the S protein from Delta is the most different from the original S protein from the Wuhan isolate (Fig. 4B). The S protein from the Delta variant presented the highest value of RMSD, 1.716 (Fig. 4D). The most altered region is the NDT region due to accumulating a higher number of mutations.

The Delta variant is more aggressive [184]. Delta-infected people present a viral load 1260 times higher than people infected with the original coronavirus strain [195,196]. Li et al. [195] reported that high viral loads in people infected with the Delta variant are reached in up to four days after contact with an infected person. In contrast, in people infected with the original SARS-CoV-2, this only happened six days after contact. This result indicates a higher replication rate of the Delta variant, which has a perfect combination to spread. The high viral loads plus short incubation time make sense to explain the greater transmissibility of the Delta variant compared to the original SARS-CoV-2 or even the other variants.

The Delta variant’s ability to strongly bind to the receptors of lung cells and escape from immune response might overcome immunization gained from vaccination [184]. Recent data indicate that fully vaccinated people can still be vectors to spread the Delta variant more efficiently than other variants [197–199]. The fully vaccinated people are protected against severe COVID-19 and death by Delta variants. However, those people still present a high level of transmission. This fact is intriguing to scientists worldwide, and the recommendation is for vaccinated people to continue taking precautions such as social distancing and wearing face masks.

10. Omicron variant

On 24 November 2021, a new variant was identified in South Africa, variant B.1.1.529, later designated as Omicron by the WHO. Two days later, Omicron was classified as a VoC, a classification that occurred faster than the other variants [200]. The Omicron variant is highly mutated, with more than 50 mutations (30 mutations in the S protein alone). There have been more than 200 cases detected in 23 countries on all 6 continents [200]. Most surprisingly, cases were detected in Scotland, where 6 cases of these variants were identified. None of them were related to travelers from South Africa, which may indicate community transmission of the variant [201].

Initial results show that this variant has many mutations in relation to the Alpha and Delta variants, linked to the sudden increase in transmissibility, infectivity, and escape from the immune system. Initial modeling studies using artificial intelligence showed that Omicron’s protein S could escape from T lymphocyte cells, but more in vitro studies need to be conducted to confirm this finding [201]. Like three patients in Brazil, some people infected by Omicron were fully vaccinated, indicating that Omicron can defeat the immunity caused by vaccines.

Preliminary studies have shown that the S protein of this new variant has an insertion at position 214 of the NTD region. This insertion was likely acquired from a host infected by SARS-CoV-2. Therefore, sequencing in the search for new variants is of paramount importance for understanding how mutations affect the behavior of protein S in human infection and the discovery of new animal hosts. That is why it is essential to continue monitoring people infected with SARS-CoV-2.

Of the 30 mutations of the S protein of the Omicron variant, some are shared with previous variants, and others are unique. Among the mutations, Omicron has one insertion of three amino acids in the NTD region together with three small deletions compared to wild-type SARS-CoV-2 from Wuhan. Insertions and deletions in the S protein from Omicron lead to a loss of three amino acid residues in the Omicron S protein compared to the wild-type version of the virus.

Indeed, mutations lead to a different S protein in the Omicron variant compared to S protein from Wuhan SARS-CoV-2 (Fig. 5). The Delta variants have the most different S protein among the variants compared to Wuhan SARS-CoV-2 (Figs. 3 and 4). However, RMSD analysis of the S protein from Omicron revealed a new winner. The S protein from the Omicron variant had an RMSD score of 2.156 (Fig. 5C), compared to the S protein structure of Wuhan isolate (Fig. 5A), with a score of 1.716 (Fig. 5B). This analysis showed many alterations in atoms’ positions, leading to a structural difference in S protein spatial arrangement. The most different parts of S protein from the Omicron variant have been found in the NTD region, which has suffered from many mutations, including insertions and deletions.

First of all, the D614G substitution in all variants is also conserved in Omicron S protein. Omicron shares mutations that were previously exclusive to Alpha, such as del69–70 and P681H. Omicron has no mutation shared only with the Beta variant, and the H655Y substitution is the only mutation that Omicron shares only with the Gamma variant. With the Delta variant, Omicron shares T95I, G142D and T478K. Omicron shares the N50IY region with Alpha, Beta, and Gamma variants. Additionally, Omicron shares the K417N substitution with Beta, Gamma, and Delta variants.

The Omicron variant has 24 signature mutations (A67V, Δ211/L212I, ins214EPE, G339D, S371L, S373P, S375F, N440K, G446S, S477N, E484A, Q493K, G496S, Q498R, Y505H, T547K, N679K, N764K, L796Y, N856K, Q954H, N969K, and L981F [201]. These mutations are essential to continue monitoring people infected with SARS-CoV-2. Preliminary studies have shown that the S protein of this new variant has an insertion at position 214 of the NTD region. This insertion was likely acquired from a host infected by SARS-CoV-2. Therefore, sequencing in the search for new variants is of paramount importance for understanding how mutations affect the behavior of protein S in human infection and the discovery of new animal hosts. That is why it is essential to continue monitoring people infected with SARS-CoV-2.

Fig. 5. Structural alignment among Spike protein from Wildtype and variants of SARS-CoV-2. (A) structural analysis of Spike protein from wildtype version of SARS-CoV-2, (B) Spike protein from Delta variant and (C) Spike protein from Omicron variant. RMSD analysis indicates changes in the structure.
Omicron.

An interesting point regarding Omicron is the substitution of E484A rather than the E484K present in the Beta and Gamma variants (Figs. 3A and 4A). As discussed above, the E484K is related to escape from neutralizing antibodies. In the Omicron variant, the E484A is more involved in escaping from 2B04 and 1B07 monoclonal antibodies [201].

As happens in other variants, the mutation in the NTD region was mostly involved in two points: stabilizing the S1 to allow strong interaction between RBD and ACE2, and escape from neutralizing antibodies. With the deletion shared with the Alpha variant, Omicron also has another del211, a substitution of L212I, and an ins214EPE of three amino acids in the NTD region. As discussed above, NDT is a supersite of recognition by antibodies. By changing this region, Omicron can be more efficient in escaping from the immune system [43,202]. These alterations could be involved in the ability of Omicron to infect fully vaccinated people.

This is not the first substitution where N440K is present. One SARS-CoV-2 sample isolated from a re-infected patient presented this mutation [151] and is related to escape from human monoclonal antibodies. This mutation is predicted to result from selective pressure imposed by antibodies [203]. Other mutations in RBD such as G446S, S477N, T478K (also in the Delta variant), E484A, and Q498R, are all involved in the escape of RBD from anti-RBD neutralizing antibodies [142,193,204].

Q493 is an important site for interaction between RBD and ACE2. However, mutations at this site are generally either Q493R/K, and the Omicron variant has the Q4 substitution. It seems that this mutation is tolerated and does not affect the interaction between ACE2 and RBD [205]. Indeed, this mutation is involved in the escape from LY-CoV555, LY-CoV016, and REGN10989/10934 monoclonal antibodies [155,205]. S477N, N501Y, P681H, and N679K are very likely to support the Omicron variant's higher transmissibility. Before Delta, Alpha was the most transmissible variant in the world. This was possible because compared to wild-type SARS-CoV-2, the Alpha variant has a histidine rather than a proline at position 681 in the polybasic cleavage site in the spike protein at the S1/S2 junction (residues 681–685) to the furin site of cleavage. The furin cleavage requires a basic environment to occur, so the replacement of proline (apolar) by histidine (positively charged) in the Alpha variant improved the furin cleavage, producing more activity of S proteins able to infect the cell. Additionally, the Alpha variant loses glycosylation at this position by removing the proline residue. This is important because carbohydrates hinder the cleavage site, inhibiting access by the protease. Elimination of proline removed a glycosylation site found in the ancestral Wuhan strain.

In the case of Delta, there has been improvement. The Delta variant has substitution at position 681, but the Delta variant has an arginine replacing the proline residue instead of lysine. By being the essential amino acid residue, the Delta variant improved the cleavage of S protein, enhancing its cell recognition and infection, leading to higher transmissibility than that presented by the Alpha variant. The Omicron preserves the same substitution presented by the Alpha variant, P681H. The replacement of proline by an arginine instead of histidine is a significant evolutionary advantage. In contrast to histidine, which is only positive at mildly basic pH, arginine is positive under all cell conditions. This is thought to be responsible for the higher transmissibility of the Delta variant [206].

At first glance it was reasonable to assume that Omicron could be less transmissible than the Delta variant. However, Omicron also showed a new mutation at the cleavage site, N679K. This substitution also added a positively charged lysine residue, making the neighborhood satisfactory for furin cleavage. So, these two mutations, P681H and N679K seem to be associated with mutations that provide a more efficient S1–S2 cleavage, leading to a more infective S protein. Although this makes total sense, it needs more investigation for confirmation. However, the rapid replacement of the Delta variant by Omicron in South Africa reinforces the hypothesis of high transmissibility rates of the Omicron variant. Another explanation for higher transmissibility of Omicron is that these mutations also promote escape from the immune system.

There is no doubt about the higher transmissibility rates of Omicron. However, it needs to be quantified. The reproductive number (R0) is a measure of transmissibility (Fig. 6). For example, the R0 of wild-type SARS-CoV-2 is 1, which means one infected person transmits the virus to one person. The R0 for the Alpha variant is 1.76, indicating more transmissibility (Fig. 6). For the Delta variant, the R0 value is 5.01 (Fig. 6), indicating that one person infected with the Delta variant can spread the virus to up to five other people. This is a high increase of transmissibility compared to wild-type SARS-CoV-2. However, there are no reliable data on R0 for Omicron (Fig. 6). Based on preliminary infection in South Africa, it is postulated to be higher than the Delta variant.

Regarding the severity of COVID-19 caused by Omicron, the knowledge is insufficient to reach any conclusions. However, there is some information available. For example, the COVID-19 caused by Omicron has new symptoms not described before regarding other variants, such as fatigue, body aches and headache. Additionally to these symptoms, one patient also presented a fever and a very high pulse rate [207,208]. Some patients have been found to be positive for Omicron infection but without symptoms [207].

Some vaccinated people have become infected with the Omicron variant, indicating this variant can to some extent, break immunity acquired with vaccination. The escape from the immune system of vaccinated people might be related to the higher number of mutations in the RBD region, allowing the escape from anti-RBD neutralizing antibodies. It has been indicated but not confirmed that Omicron can escape from memory immunity conferred by T cells and durable immunity.

Among vaccinated people infected with Omicron, some are asymptomatic and some have only mild symptoms. There have been no severe cases, hospitalization, or death caused by Omicron. This indicates that vaccination does not prevent Omicron infection, but it protects from severe cases, hospitalization, and death [200,201,207,208]. This information suggests how important it is to expand vaccination worldwide to prevent severe COVID-19 caused by the Omicron variant.

11. Interaction of S protein from variants with ACE2

Although it is known that the S protein from variants has a higher affinity for ACE2 than the wild-type version, nobody has quantified this interaction. What is known and accepted is that variants are more transmissible than wild-type versions. Here, we have used the wild-type S protein from the Protein Data Bank (https://www.rcsb.org/, PDB ID: 6Z97) to build the three-dimensional (3D) models of protein from the Alpha, Beta, Gamma, Delta and Omicron variants using the Swiss Model Server (https://swissmodel.expasy.org/) [209]. All the structure showed in the manuscript were validated. The models were created using the SWISS-model that already make the check in the structure using the Molprobity. Additionally, we have checked the structures on Molprobity server (http://molprobity.biochem.duke.edu/) [210] by Ramachandran plot analysis. Also, the global quality factor and the reliability of the models' folding were evaluated by ERRAT2 (http://services.mbi.ucla.edu/ERRAT/) and Verify3D (http://services.mbi.ucla.edu/Verify3D/), respectively. The 3D model built was docked against the ACE2 (PDB ID: 7KMD) using the FRODOCK Interactive protein-protein docking server (http://frodock.chaconlab.org/) [211]. The FRODOCK server provides a docking score (DS) that indicates the interaction between the protein and ligand. In this case, higher DS values are associated with higher affinity of the interaction. Based on the corresponding DS value, this only enables suggesting which protein has more affinity than any other protein. As a control, the interaction between the wild-type S protein from Wuhan isolate with ACE2 (Fig. 7A) has been used. This interaction has a DS value of 3025.00 (Fig. 7A). All S proteins from variants have higher DS values, indicating they have a higher affinity for ACE2 than wild-type S protein (Fig. 7).
expected, because, to some extent, all variants are more transmissible than wild-type SARS-CoV-2 from Wuhan. The S proteins from Beta and Gamma variants presented DS values, respectively, of 3102.00 and 3200.00, which mean increases of 2.5% and 5.85% in the DS interaction with ACE2 (Fig. 7C and D). This result suggests the S protein from those variants has a slight increase in the ability to interact with ACE2 compared to the wild-type versions of S protein.

The DS values for Alpha, Delta, and Omicron were, respectively, 3341.00, 3415.00, and 4150.00 representing increases of 10.44%, 12.89%, and 37.23% in the ability to bind to ACE2 (Fig. 7B, E, and F). These three variants were classified as transmissible and presented the highest DS values compared to other wild-type variants of S protein. Not surprisingly, the S protein from the Omicron variant, which has many mutations, presented the highest value of DS, suggesting the S protein from the Omicron variant is the most optimized protein for interaction with the ACE2 receptor. These results are in accordance with the higher transmissibility of the Omicron variant.

12. Conclusion

It is clear that the S protein from variants has been changing with two goals: 1) to improve the interaction with ACE2; and 2) to escape from the immune system. Spike amino acid substitutions, deletions, and insertions (Omicron) have a direct impact on how neutralizing antibodies interact with S protein from variants, leading to reinfection and even infection of fully vaccinated people. However, it looks like the mutations in the spike of variants are more driven to improve the interaction with the ACE. The fast interaction with ACE2 reduces time when SARS-CoV-2 is exposed to neutralizing antibodies. The results presented lead to the conclusion that the Omicron S protein is optimized to bind to the ACE2 receptor and escape from the immune system. It is clear that SARS-CoV-2 is under selective pressure caused by vaccines and is changing in response to that pressure. So, an understanding of spike mutations in antigenic positions will help monitor and improve vaccine effectiveness. Additionally, it is essential to continue collecting virus information by sequencing and studies with vaccines to elucidate the SARS-CoV-2 behavior.

Funding and acknowledgments.

Special thanks go to CAPES for providing the postdoctoral grant (grant number 88887.318820/2019-00) to Pedro F. N. Souza. This work was supported by grants from the following Brazilian agencies: National Council for Scientific and Technological Development (CNPq), research productivity grant to Raquel C. Montenegro; Office to Coordinate Improvement of Higher Education Personnel (CAPES) for the grant to Raquel C. Montenegro (number 88881.505364/2020-01 - Call No. 9/2020; National Institute for Science and Technology of Bioinspiration and Cearense Foundation to Support Scientific and Technological Development (FUNCAP), for the grant to Raquel C. Montenegro (Number: 03195011/2020). Raquel C. Montenegro and Débora S.C.M. Castelo-Branco acknowledge the School of Medicine (FAMED) of Federal University of Ceará for supporting this research by the COVID-19 research program.

CRediT authorship contribution statement

Conceptualization: Pedro F. N. Souza.
Data curation: Pedro F.N. Souza, Felipe P. Mesquita, Jackson L. Amaral, Patrícia G. C. Landim, Karollyny R.P. Lima, Marília B. Costa, Izabelle R. Farias, Mônica O. Belem, Yago O. Pinto, Heline H. T. Moreira, Ilana C.L. Magalhaes, Claudia R. de Andrade.

Formal analysis: Pedro F.N. Souza, Felipe P. Mesquita, Jackson L. Amaral, Patrícia G. C. Landim, Karollyny R.P. Lima, Marília B. Costa, Izabelle R. Farias, Mônica O. Belem, Yago O. Pinto, Heline H. T. Moreira, Ilana C.L. Magalhaes, Claudia R. de Andrade.

Funding acquisition: Pedro F.N. Souza.

Resources: Pedro F.N. Souza.

Supervision: Pedro F.N. Souza.

Writing–original draft: Pedro F.N. Souza, Patrícia G. C. Landim, and Claudia R. de Andrade.

Writing, review, and editing: Pedro F.N. Souza.

Final approval of the manuscript and submission: Pedro F.N. Souza.

Ethical approval

Not applicable.

Declaration of competing interest

All authors declare no conflicts of interest.

References

[1] S.R. Weiss, Forty years with coronaviruses, J. Exp. Med. 217 (2020), https://doi.org/10.1084/jem.20200537.

[2] Coronavirus Disease (COVID-19) situation reports, (n.d.). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports (accessed January 21, 2021).

[3] Coronavirus disease (COVID-19), (n.d.). https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed January 21, 2021).

[4] Coronavirus update (live): 97,633,202 cases and 2,091,064 deaths from COVID-19 virus pandemic - Worldometer, (n.d.). https://www.worldometers.info/coronavirus/ (accessed January 21, 2021).

[5] WHO Coronavirus Disease (COVID-19) dashboard | WHO Coronavirus Disease (COVID-19) dashboard, (n.d.). https://covid19.who.int/ (accessed January 22, 2021).

[6] C. Drosten, B. Meyer, M.A. Mülher, V.M. Gorman, M. Al-Masri, R. Hossain, H. Madani, A. Sieberg, B.J. Bosch, E. Lattwein, R.F. Alhakeem, A.M. Ansari, W. Hajomar, A.M. Albarak, J.A. Al-Tawfiq, A.I. Zama, Z.A. Memish, Transmission of MERS-coronavirus in household contacts, N. Engl. J. Med. 371 (2014) 828–835, https://doi.org/10.1056/nejmoa1405858.

[7] A.M. Zaki, S. van Boheemen, T.M. Bestebroer, A.D.M.E. Osterhaus, R.A. M. Fouchier, Isolation of a novel coronavirus from a man with pneumonia in
immune cells in patients with COVID-19, Nat. Med. 26 (2020) 842–844, https://doi.org/10.1038/s41591-020-0991-0.

[33] Y. Yuen, S. Alvaldah, M. Shi, G. Godman, M. Zuo, J.A. Madison, C. Blair, A. Weber, B.J. Barnes, M. Egeblad, R.J. Woods, Y. Kanthi, J.S. Knight, Neutrophil extracellular traps in COVID-19, JCI Insights 5 (2020), https://doi.org/10.1172/jci.insight.138993.

[34] S. Berens, R. Dijkman, M. Habjan, A. Heurisch, S. Gierer, I. Glowacka, K. Welteh, M. Winkler, H. Schneider, H. Hoffmann-Winkler, V. Thiel, S. Pohlmann, TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium, J. Virol. 87 (2013) 6150–6160, https://doi.org/10.1128/JVI.0372-12.

[35] L. Du, R.Y. Kao, Y. Zou, Y. He, G. Zhao, C. Wang, S. Jiang, K.Y. Yen, D.Y.-J. Yin, B.-J. Zheng, Cleavage of spike protein of SARS coronavirus by protease factor x is associated with viral infectivity, Biochem. Biophys. Res. Commun. 359 (2007) 174–179, https://doi.org/10.1016/j.bbrc.2007.06.092.

[36] O. Krocklin, Y. Li, A. Andonov, H. Feldmann, R. Flick, S. Jones, U. Stroehler, N. Bastien, K.V.N. Daurut, K. Cheng, J.N. Simonsson, H. Perreault, J. Wilkins, W. Eise, F. Plummer, K.G. Standing, Mass spectrometric characterization of proteins from the SARS virus: a preliminary report, Mol. Cell Proteomics 2 (2003) 346–356, https://doi.org/10.1021/bi026615p.

[37] B. Delman, H. Laude, Assembly of coronavirus spike protein into trimers and its role in epitope expression, J. Virol. 89 (2015) 5367–5375, https://doi.org/10.1128/JVI.01567-15.

[38] Y. Watanabe, J.D. Allen, D. Wrapp, J.S. McLellan, M. Crispin, Site-specific glycan analysis of the SARS-CoV-2 spike, Science 369 (2020) 330–333, https://doi.org/10.1126/science.abb2124.

[39] H.C. Song, M.-Y.seo, K. Stadler, B.J. Yoo, Q.-L. Choo, S.R. Coates, Y. Uematsu, T. Harada, C.E. Greer, J.M. Polo, P. Pileri, M. Eckmann, R. Rappuoli, S. Abrignani, M. Houghton, J.H. Han, Synthesis and characterization of a native, enzymatic form of recombinant human coronavirus 229E spike glycoprotein, J. Virol. 78 (2004) 10328–10335, https://doi.org/10.1128/JVI.78.22.10328-10335.2004.

[40] D. Wrapp, N. Wang, K.S. Curbert, J.A. Goldsmith, C.L. Hsieh, O. Abiona, B. S. Graham, J.S. McLellan, Structural and functional basis of SARS-CoV-2 entry by using human ACE2, Cell 181 (2020) 894–904, https://doi.org/10.1016/j.cell.2020.02.045, e9.

[41] J.K. Miller, G.R. Whittaker, Physiological and molecular triggers for SARS-CoV-2 membrane fusion and entry into host cells, Virilocity 517 (2018) 3–8, https://doi.org/10.1126/science.abb2507.

[42] Z.A. Memish, S. Perlman, M.D. Van Kerkhove, A. Zulima, Middle East respiratory syndrome, Lancet 395 (2020) 1063–1073, https://doi.org/10.1016/S0140-6736(20)30522-3.

[43] S. Liu, G. Xiao, Y. Chen, Y. He, J. Niu, C.R. Escalante, H. Xiong, J. Farmar, A. Debnath, P. Tien, S. Jiang, Interaction between heptad repeat 1 and 2 regions of SARS-CoV-2 spike (S) protein originating in China, Cell Host Microbe 27 (2020) 325–328, https://doi.org/10.1016/j.chom.2020.02.001.

[44] M.R. Alam, M.R. Kahler, S. Roza, Comorbidities might be a risk factor for the incidence of COVID-19: evidence from a web-based survey, Prev. Med. Reports. 21 (2021), 101319, https://doi.org/10.1016/j.pmedr.2021.101319.

[45] N. Zhu, W. Wang, Z. Liu, C. Liang, W. Wang, F. Ye, B. Huang, L. Zhao, H. Wang, W. Yu, Z. Deng, L. Mao, C. Su, G. Gao, T. Jiang, J. Zhao, G. Wu, J. Song, W. Tan, Morphogenesis and cytopathic effect of SARS-CoV-2 infection in human airway epithelial cells, Nat. Commun. 11 (2020) 1–8, https://doi.org/10.1038/s41467-020-17796-2.

[46] S. Hoffmann, M. Kleiner-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Eichsen, T. Schiergerns, G. Herrler, N.H. Wu, A. Nitsche, M.A. Müller, C. Drosten, Single-cell landscape of bronchoalveolar–lavage–cell–trials, J. Biomol. Struct. Dyn. (2021) 1–10, https://doi.org/10.1080/07391102.2020.1871415.

[47] A.C. Walls, J.V. Park, M.A. Tortora, A. Wall, A.T. McGuire, D. Veesler, Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein, Curr. Opin. Microbiol. 51 (2020) 281–292, https://doi.org/10.1016/j.copmi.2020.02.058, ebo.

[48] Z. Xu, L. Shi, Y. Zhang, J. Huang, C. Zhang, S. Liu, P. Zhao, H. Liu, L. Zhu, Y. Tai, C. Bai, T. Gao, J. Song, P. Xia, J. Dong, J. Zhao, F.S. Wang, Pathological findings of COVID-19 associated with acute respiratory distress syndrome, Lancet Respir. Med. 8 (2020) 422–423, https://doi.org/10.1016/S2213-2600(20)30076-X.

[49] D. Wang, B. Hu, C. Fu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, X. Xiong, Y. Zhao, Y. Li, X. Wang, Z. Peng, Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China, JAMA. J. Am. Med. Assoc. 323 (2020) 1061–1069, https://doi.org/10.1001/jama.2020.1585.

[50] M. Dong, J. Zhang, X. Ma, J. Tan, L. Chen, S. Liu, Y. Xin, I. Zhang, ACE2, TMPRSS2 distribution and extrapulmonary organ injury in patients with COVID-19, Biomed. Pharmacother. 131 (2021), 110678, https://doi.org/10.1016/j.biopha.2020.110678.

[51] V.G. Puelles, M. Lütgehetmann, M.T. Lindenmeyer, J.M. Schrepfer, N.M. Wong, L. Alwiss, S. Chilla, A. Heinemann, N. Wanner, S. Liu, F. Braun, S. Lu, S. Feilser, A.S. Schöttler, C. Edler, O. Grozinger, T. Wich, S. Kluge, P. Kuespel, M. Felczerbauer, T.B. Hüber, Multiforgan and renal tropism of SARS-CoV-2, N. Engl. J. Med. 383 (2020) 590–592, https://doi.org/10.1056/NEJMoa2003766.

[52] Y. Tang, J. Liu, D. Zhang, Z. Xu, J. Ji, C. Wen, Cytokine storm in COVID-19: the current evidence and treatment strategies, Front. Immunol. 11 (2020) 1708, https://doi.org/10.3389/fimmu.2020.01708.

[53] Z. Varga, A. Flamant, F. Stegmann, M. Habeler, R. Andernatt, A. S. Zinkernagel, M.R. Mehra, R.A. Schupbach, F. Ruschitska, H. Moch, Endothelial cell infection and endothelitis in COVID-19, Lancet Respir. Med. 3 (2020) 1417–1418, https://doi.org/10.1016/S2213-2600(20)30037-5.

[54] Y. Tang, J. Liu, Y. Huang, Y. Cai, G. Xu, J. Qian, L. Cheng, J. J. X. Wang, F. Wang, L. Liu, I. Amlie, Z. Zhang, Spike-cell–like–state of bovine coronaviral
P.F.N. Souza et al.

H. Fort, A Very Simple Model to Account for the Rapid Rise of the British Variant J. Hu, P. Peng, K. Wang, L. Fang, F. Yang Luo, A. Shun Jin, B. Zhong Liu, N. Tang, C. Rees-Spear, L. Muir, S. Griffith, J. Heaney, Y. Aldon, J. Snitselaar, P. Thomas, J. Edgeworth, J. Breuer, J.A. Todd, C. Fraser, D. Buck, M. John, G.L. Kay, R. Young, C.M. McCann, S. Elliott, H. Lowe, B. Temperton, S. Roy, A. Price, R. Dewar, A.S. Jahun, C. McMurray, S. Pandey, J.P. McKenna, A. Nelson, G. M. Bashton, A.P. Underwood, N. Pacchiarini, K.F. Loveson, K.E. Templeton, C. S. Moses, R. Amato, S. Nicholls, M. Bull, D.L. Smith, J. Barrett, D.M. Aanensen, M. E.M. Harrison, S. Goncalves, D.J. Fairley, M.W. Loose, J. Watkins, R. Livett, S. Lytras, A. Mohamed, G. Gallo, N. Thakur, D.A. Collier, P. Mlcochova, S. Neutralization, BioRxiv, 2021, https://doi.org/10.1101/2021.01.15.426849, R. Gupta, P. Cherepanov, H. Stauss, E. Nastouli, , on behalf of the S. Investigators, 10.1016/J.CELL.2021.03.028/ATTACHMENT/F72F3A1A-8DD6-4517-A111-R. Abdelnabi, S.Y.C. Foo, P.W. Rothlauf, L.M. Bloyet, F. Benigni, E. Cameroni, H69/V70, BioRxiv, 2021, https://doi.org/10.1101/2020.12.14.422555, K. Gupta, Recurrent Emergence and Transmission of a SARS-CoV-2 Spike Deletion 10.1002/VETR.944 no-no. and cats with suspected myocarditis, Vet. Rec. 189 (2021), https://doi.org/10.3390/V13091759, 13 (2021) 1759. refVal

B. Farr, M. Fenton, M. Ferrero, N. Flack, H. Fordham, G. Forsythe, M. Francis, A. Dove, L. Drummond, M. Dudek, L. Durrant, E. Easthope, S. Eckert, P. Ellis, J. S. Green, L. Harris, M. Hendrikx, L. Hoang, T. I. Nordgren, D. J. Rodríguez, M. Monyou, D. M. Murray, J. S. Njue, A. O. Toole, A.R. Hesketh, R. Stark, L. du

A. O. Toole, A.R. Hesketh, R. Stark, L. du

MacIntyre-Cockett, A. Birchley, A. Adams, A. Plimmer, B. Gatica-A. Thornton, R. Myers, J. Stockton, C.A. Williams, W.C. Yew, A.J. Trotter, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O’Toole, J. Watts, C. Breen, K. A. Jackson, S.S.Y. Ahma...
[191] G. Lu, Y. Hu, Q. Wang, J. Qi, F. Gao, Y. Li, Y. Zhang, W. Zhang, Y. Yuan, J. Bao, B. Zhang, Y. Shi, J. Yan, G.F. Gao, Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26, Nature 500 (2013) 227–231, https://doi.org/10.1038/nature12328.

[192] C.O. Barnes, C.A. Jette, M.E. Abernathy, K.M.A. Dam, S.R. Esswein, H.B. Gristick, A.G. Malyatin, N.G. Shafar, K.E. Huay-Tubman, Y.E. Lee, D.F. Robbiani, M.C. Nissenweerg, A.P. West, P.J. Bjorkman, SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies, Nat. 5887839 (588) (2020) 682–687, https://doi.org/10.1038/s41586-020-2852-1, 2020.

[193] Y. Liu, J. Liu, B.A. Johnson, H. Xia, Z. Ku, C. Schindewolf, S.G. Widen, Z. An, S.C. O. Barnes, C.A. Jette, M.E. Abernathy, K.M.A. Dam, S.R. Esswein, H.B. Gristick, J.A. Plante, B.M. Mitchell, K.S. Plante, K. Debbink, S.C. Weaver, V.D. Menachery, B. Li, A. Deng, K. Li, Y. Hu, Z. Li, G. Haidar, W.Paul Duprex, Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape, Science 371 (2021) 1139–1142, https://doi.org/10.1126/science.abf9950.

[194] P.R. Ranu, M. Imran, J.V. Lakshmi, B. Jolly, A. Jain, A. Surekha, V. Senthivel, P. Chandrasekhar, M.K. Divakar, D. Sinivasulu, R.C. Bhatia, P.R. Vanaja, V. Scaria, S. Sivasubbu, Symptomatic reinfection of SARS-CoV-2 with spike protein variant N440K associated with immune escape, J. Med. Virol. 93 (2021) 4163–4165, https://doi.org/10.1002/jmv.26997.

[195] P. Arora, C. Rocha, A. Kempf, I. Nehmleier, L. Graichen, M.S. Winkler, M. Lier, S. Schulz, H.-M. Jack, A. Cosmann, M.V. Stankov, G.M.N. Behrens, S. Pohlmann, M. Hoffmann, The spike protein of SARS-CoV-2 variant a.30 is heavily mutated and evades vaccine-induced antibodies with high efficiency, Cell. Mol. Immunol. 18 (2021) 2673–2675, https://doi.org/10.1038/s41423-021-00779-5.

[196] T.N. Starr, A.J. Greaney, A.S. Dingens, J.D. Bloom, Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016, Cell Rep. Med. 2 (2021), https://doi.org/10.1016/J.CHRM.2021.100255/ATTACHMENT/ESFBD729-E3F2-4C56-8D37-48D273794D59-MMC2-CSV.

[197] L. Zhang, M. Mann, Z.A. Syed, H.M. Reynolds, E. Tian, N.L. Samara, D.C. Zeldin, L.A. Tabak, K.G. Ten Hagen, Furin cleavage of the SARS-CoV-2 spike is modulated by O-glycosylation, Proc. Natl. Acad. Sci. 118 (2021), e2109905118, https://doi.org/10.1073/PNAS.2109905118.

[198] How bad is Omicron? What scientists know so far, (n.d.), https://www.nature.com/articles/d41586-021-03614-z (accessed December 8, 2021).

[199] How is Omicron different from Delta? What scientists say, (n.d.), https://www.nature.com/articles/d41586-021-03614-z (accessed December 8, 2021).

[200] Classification of Omicron (B.1.1.529): SARS-CoV-2 variant of concern, (n.d.), https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern (accessed December 8, 2021).

[201] E. Callaway, Heavily mutated coronavirus variant puts scientists on alert, Nature (2021), https://doi.org/10.1038/D41586-021-03552-W.

[202] R.K. McCarthy, L.J. Remick, S. Nambali, L.R. Robinson-McCarthy, W.G. Bain, G. Haidar, W. Paul Duprex, Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape, Science 371 (2021) 1139–1142, https://doi.org/10.1126/science.abf9950.

[203] Omicronvariant border bans ignore the evidence, say scientists, (n.d.), https://www.nature.com/articles/d41586-021-03614-z (accessed December 8, 2021).

[204] What is the real impact of Omicron? Scientists give their views, (n.d.), https://www.nature.com/articles/d41586-021-03614-z (accessed December 8, 2021).

[205] Levine, FRODOCK 2.0: fast protein–protein docking server, Bioinformatics 32 (2016) 2386–2388, https://doi.org/10.1093/BIOINFORMATICS/BTW141 (accessed December 8, 2021).