© Indian Academy of Sciences

Review

Why are some coronavirus variants more infectious?

RAJU MUKHERJEE* and ROHIT SATARDEKAR

Department of Biology, Indian Institute of Science Education and Research (IISER) Tirupati, Tirupati, India

*Corresponding author (Email, raju.mukherjee@iisertirupati.ac.in)

Since the start of the pandemic, SARS-CoV-2 has infected almost 200 million human hosts and is set to encounter and gain entry in many more in the coming months. As the coronavirus flourish, the evolutionary pressure selects those variants that can complete the infection cycle faster and reproduce in large numbers compared to others. This increase in infectivity and transmissibility coupled with the immune response from high viral load may cause moderate to severe disease. Whether this leads to enhanced virulence in the prevalent Alpha and Delta variants is still not clear. This review describes the different types of SARS-CoV-2 variants that are now prevalent, their emergence, the mutations responsible for their growth advantages, and how they affect vaccine efficacy and increase chances of reinfection. Finally, we have also summarized the efforts made to recognize and predict the mutations, which can cause immune escape and track their emergence through impactful genomic surveillance.

Keywords.   SARS-CoV-2; variants of concern; variants of interest; infectious; immune escape

All variants are innocent unless proven guilty.

— Eric Topol, The Scripps Research Institute, California.

1. Introduction

It is close to 2 years since the COVID-19 pandemic started, but reports of new infections from the SARS-CoV-2 are still on the rise. While multiple vaccines of different types are now available, the cause of concern is complete protection from being infected with the new antigenic variants of the original strain first detected in Wuhan, China. Dorothy Hamre first isolated the human coronavirus from the respiratory tract of students admitted to the infirmary at the University of Chicago in 1966 (Hamre and Procknow 1966). This common cold causing strain was later named 229E, which along with OC43, the strain believed to be the reason for the pandemic in cattle and humans in 1890 (King 2020), has been widely prevalent in the world, and most of us probably carry antibodies against them.

Coronavirus belongs to the order of Nidovirus, which are enveloped positive sense single strand RNA viruses that generate $3'$ co-terminal nested sub-genomic mRNAs to code for its structural and accessory proteins and mostly infect animals (Di *et al*. 2018). Being an RNA virus, they have a tremendous ability to mutate as they replicate within their animal reservoir or after they jump to a new species in order to adapt to a new host. A change in only two amino acids was enough for SARS-CoV-1 to adapt in human cells (Li *et al*. 2005). Similarly, it is believed that a single amino acid change (T372A) may have helped SARS-CoV-2 to spillover from bats to humans (Kang *et al*. 2021). To make this emergence a success, the virus must evolve fast to gain in its ability to subvert the human immune system and

replicate inside the human cells, and this process might happen through several rounds of infection and reinfection. The more the genome replicates, the more the number of mutations that get accumulated as a result, several thousand of variants are produced as it evolves.

This review describes the different types of SARS-CoV-2 variants that are now prevalent, how they are formed, how it affects their infectivity, COVID-19 disease severity, and transmission. More importantly, how do their presence and emergence impact the effectiveness of the current vaccines and future breakthrough infections?

2. What is a variant?

RNA viruses generally have a very high error rate in replication, which on the one hand, is beneficial for them to adapt to selective pressures in the environment but generates a large number of self-eliminating deleterious mutants on the other. This feature also makes developing any chemotherapeutic or vaccine intervention challenging. The low replication fidelity of the RNA-dependent RNA polymerase is partially compensated by the $3'–5'$ exonuclease proofreading activity present in the non-structural protein (nsp14) (Robson *et al.* 2020). This has resulted in SARS-CoV-2 accumulating mutation at a slower rate of around one in every two weeks, corresponding to 1.1×10^{-3} substitutions per site per year. Until October 2020, most genomes were found to collect around twenty mutations compared to the ancestral strain (Wuhan-Hu-1), suggesting that the virus was constantly changing its genome while replicating and sampling the fitness landscape inside a wide range of immunocompetent humans, its new host. Each time the ancestral sequence is changed, a new variant is born. Few survive while most disappear from the population. The mutations occur randomly distributed across the genome, and the ones that were associated with low-fitness cost or no-cost were tolerated and remained unnoticed. However, a small number of them did give the coronavirus a growth advantage by altering its infection cycle by increasing its infectivity and transmissibility. Importantly, some regions in the genome become less conserved and can tolerate diversity, thus representing the hotspots for generating new variants. In an epidemic, these changes are monitored in a community through genomic surveillance to track the ancestral strain and build the phylogenetic tree in which each persisting branch becomes a lineage. At the start of the COVID-19 pandemic, two clusters comprising strain A (20%)

and B (80%) emerged in China, while the progenitor strain A slowly disappeared, strain B became the most dominant and mutated to generate more lineages as it spread outside China (Rambaut *et al.* 2020). The first report of a non-synonymous mutation leading to an amino acid change from aspartic acid to glycine at position 614 in the most important spike glycoprotein (aka D614G) came in from Scotland in March 2020 (Robertson 2020). This strain was later christened as B.1, was responsible for the biggest outbreak in Lombardy, Italy. The strain B.1 appeared to gain a selective advantage over others in increasing the magnitude of infection by producing more viral load. However, this strain had no impact on the disease severity and went on to produce B.1.1–B.1.1.7 in the United Kingdom (Rambaut *et al.* 2020). Since then, several national genomic epidemiology efforts have sequenced thousands of isolates and curated a database (GISAID) of the genetic changes associated with major epidemiological events and changes in the virus's behavior (Elbe and Buckland-Merrett 2017). The GISAID dashboard helped track emerging lineages as they traveled between regions and countries and captured the global pattern of the SARS-CoV-2 genetic diversity (figure 1).

3. How many SARS-CoV-2 variants are there?

As the virus spread across the continents, it replicated unfettered and generated as many mutants as possible to produce a large number of variants, which differs from each other by the presence of a set of distinctive mutations in its genome. The Pango lineage database, which analyzes the genome sequences deposited in GISAID, has classified the genomes into 1277 variants or lineages (O'Toole and Pybus 2021). These mutations are spread across the whole genome of the coronavirus, but not all of the changes have contributed in altering the viral physiology. For example, the B.1.525 (Eta) variant first detected from Nigeria reported the maximum number of mutations in the genome but with no proportional changes in the viral behavior (Haseltine 2021). Some mutations have caused a significant change in the shape or number of the viral surface proteins, which have resulted in either increasing the virus's infectivity and transmissibility or its ability to cause severe disease. In addition, these mutations may cause interference in diagnostics and more importantly, may increase chances of reinfection and bring down vaccine efficacy if the modified viral proteins are able to escape from being recognized by
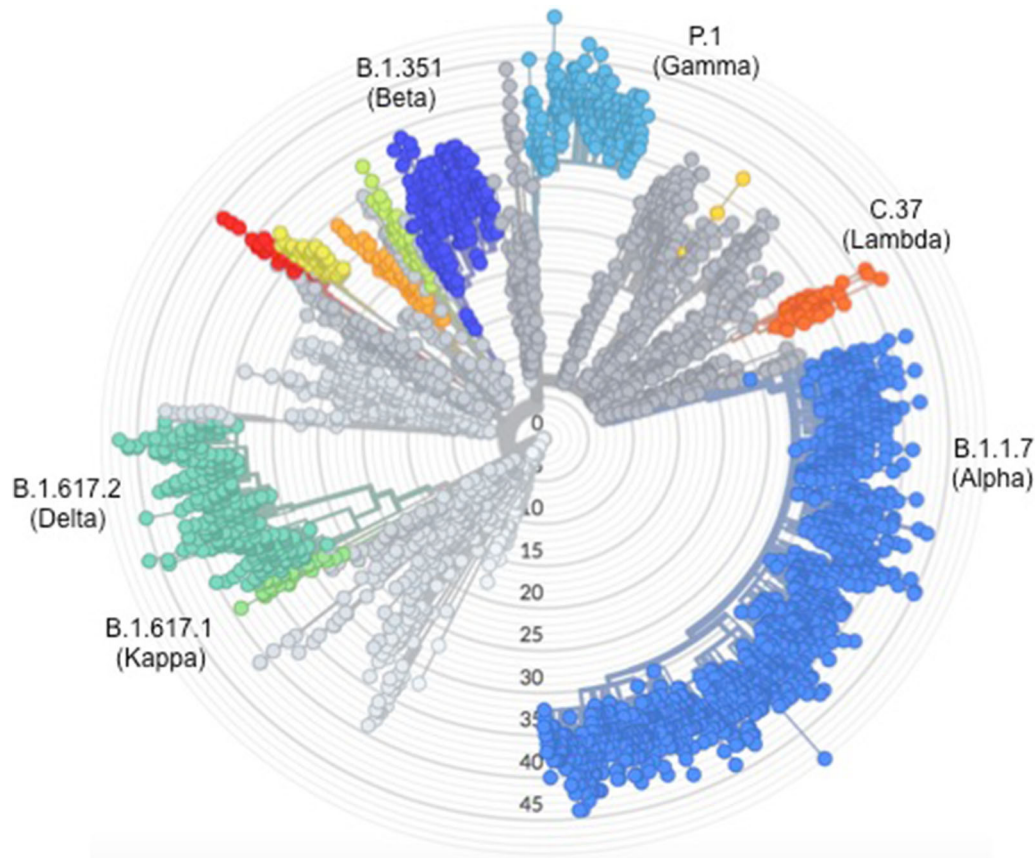
Figure 1. Genomic epidemiology of the emerging lineages of SARS-CoV-2: phylogenetic divergence of 3869 strains sequenced between December 2020 and July 2021. Figure created from *https://nextstrain.org/* and colored as per clades.

the neutralizing antibody generated from previous infections and vaccines. Several dominant strains have become prevalent by natural section and have given away to newer variants over the course of the pandemic (figure 2; table 1). However, these beneficial mutations have become worrisome for us as they get reported from large cluster outbreaks and raise concerns of the protection provided by the currently available vaccines.
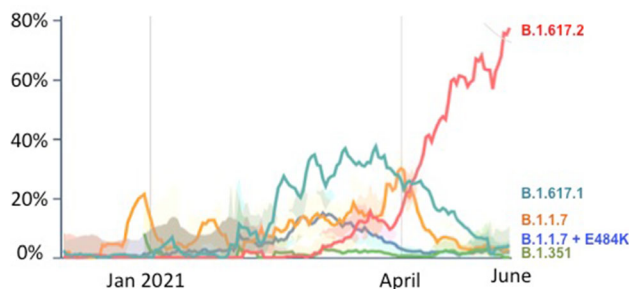

Figure 2. Dominance of the B.1.617.2 (Delta) variant: prevalence of SARS-CoV-2 lineages over time in India. Line plot for seven day average and shaded region corresponding to the 95% CI. Created from *https://outbreak.info* server.

The D614G strain (B.1) that emerged in China became the first 'variant of concern' as its prevalence increased from 10% in March 2020 to 78% of all the sequences analyzed by May (Robertson 2020). This D to G change was always accompanied by three other mutations, including an amino acid change in the RNA-dependent RNA polymerase. When checked for its infectivity, pseudovirus expressing the D614G modified spike protein produced nine fold more infectious titres, justifying the higher magnitude of infection it caused in the real world (Korber *et al.* 2020). Until September 2020, the SARS-CoV-2 genome accumulated ten mutations, when the B.1.1.7 (Alpha) strain with 23 mutations (8 in the spike protein) was detected in Kent, United Kingdom, and was reported to be 50% more transmissible (Rambaut *et al.* 2020). However, it did not show any reduction in vaccine efficacy. Like its ancestor B.1 strain, it also produced a high viral load leading to a low Ct value in the RT-PCR-based diagnosis (Collier *et al.* 2021). It is hypothesized that B.1.1.7 may have evolved in patients with weak immunity who received convalescent

Table 1. Effect of the different mutations on transmission and vaccine efficacy in SARS-CoV-2 lineages.

Variants	Country of origin	Mutations in spike protein	Infectivity	hACE2 binding	Immune escape	Vaccine efficacy
B.1.1.7 (Alpha)	United Kingdom	**N501Y, ΔH69/ ΔV70,** ΔY144, A570D, D614G, T716I, S982A, **P681H**	60% higher transmission	2–5 fold higher	Minimal reduction in neutralization	Pfizer-89.5% Oxford–AZ-75-84% Novavax-86-96%
B.1.351 (Beta)	South Africa	**K417N, E484K, N501Y** D80A, D215G, ΔL242, ΔA243, ΔL244, D614G	50% higher transmission	5 fold higher	6-7 fold reduced neutralization by human CP and RBD, NTD targeting mAbs	Pfizer-75.0% Oxford–AZ-10% J&J-60%
B.1.427/ 429 (Epsilon)	USA (California)	**L452R,** S13I, W152C, D614G	20% higher transmission	2-3 fold higher	Reduced neutralization by human CP and RBD, NTD targeting mAbs	Pfizer-89.5% Oxford-AZ-75%
P.1 (Gamma)	Brazil	**K417T, E484K, N501Y,** L18F, T20N, P26S, D138Y, R190S, T1027I, H655Y	Higher transmission	2–3 fold higher	Reduced neutralization by human CP	J&J-68%
B.1.617.2 (Delta)	India	T19R, G142D, Δ156/Δ157, R158G, **L452R, T478K,** D614G, **P681R,** D950N	60% more transmissible than B.1.1.7	No significant change in affinity	3-6 fold reduction in neutralization by vaccine sera and human CP	Pfizer-88% Oxford-AZ-64% Bharat Biotech-65%

CP, convalescent plasma; AZ, AstraZeneca; J&J, Johnson & Johnson.

Dominant mutations recognized to alter infectivity or binding to antibodies are marked in bold

Source: *www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html*.

plasma therapy containing antibodies, contributing to selecting this variant (figure 3).

Similarly, another lineage of the B.1 variant, B.1.351 (Beta), evolved in South Africa and became a dominant strain in the region with 47% of all the sequences analyzed (Tegally *et al.* 2021). Its genome showed two additional mutations, which allowed its spike protein to bind tightly to the ACE receptor and at the same time evade binding to the neutralizing antibodies present in convalescent plasma and that were produced by the Astrazeneca vaccines by around a factor of ten (Wang *et al.* 2021). Monoclonal antibodies formulated by Eli Lilly and one present in the Regeneron cocktail also showed markedly reduced activity against the mutated spike protein, 'Eek mutation (E484K)' present in this variant (Annavajhala *et al.* 2021).

Among all variants of concern, the super-infectious B.1.617, which along with the Alpha variant was responsible for the collapse of the Indian health care system, has now become the most dominant lineage in India and 80 other countries (Yadav *et al.* 2021). This strain has 13 mutations, including two new in the spike protein (E484Q and P681R), attributed to its 50% higher transmissibility compared to the alpha variant. Since then, three subtypes B.1.617.1/2/3 have emerged and over time, the B.1.617.2 (Delta) have replaced the other two less transmissible lineages to become the dominant variant (Bolze *et al.* 2021). Despite lacking mutations at position N501 and E484 in its spike protein, this variant has also been found to spread faster within the body and *in vitro* was found to be less sensitive to the BNT162b2 (BioNTech/Pfizer) vaccine (Collier *et al.* 2021; Wall *et al.* 2021). These increased features of ACE2 binding and S1/S2 cleavage have been attributed to the two new mutations at positions L452R and E484Q in the receptor binding domain of the delta variant (Cherian *et al.* 2021).

While in South America, another variant with 17 mutations became predominant. This strain P.1 (Gamma), also evolved from B.1.1.28 lineage and
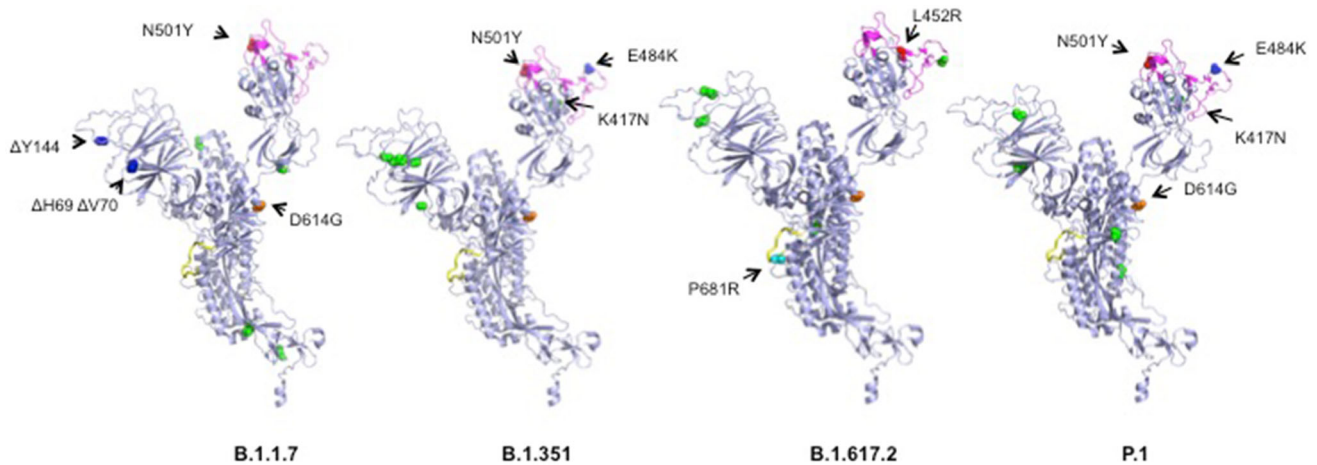
Figure 3. Mutations in Spike protein S1 subunit have implications in increasing infectivity and reducing vaccine efficacy in COVID-19. Mutations which increase the affinity of the receptor binding domain to the human ACE2 receptor (red) and help escape binding to neutralizing antibody (blue). D614G substitution (orange) stabilizes prefusion conformation and contributes in tighter binding to ACE2. Mutations responsible for increased viral fusion (cyan) by increasing cleavage at the furin cleavage site (yellow). Receptor binding motif at the CTD of the S1 subunit in magenta while all other mutations are depicted in green.

shared the three important mutations present in the receptor binding domain of the Beta variant that are associated with high transmissibility and ability to re-infect (Faria *et al*. 2021a). The Gamma variant reached the United States around the same time as another variant B.1.429, which evolved and dominated the state of California before it declined to become a *'variant of interest'* (WHO 2021b). Several other strains (B.1.526, B1.617.1, P.3), which showed higher affinity for receptor binding and potential for immune escape *in vitro* due to the presence of the same mutations have been classified as variants of interest but have not been reported to be more contagious. Intriguingly, the B.1.526 variant was able to cause reinfection in New York in the face of active immunity developed from the first 2020 wave. Therefore, it becomes essential to investigate how these mutations are being formed. Are they present in the immune suppressed population, or whether they develop from persistent infections? Monitoring of new variants such as the B.1.617.3 generated in a region and assessing their phenotypic properties must continue in order to evaluate vaccine efficiency.

4. SARS-CoV-2 infection cycle

To understand how each of these genetic changes can impact the spread of COVID-19, it is important to know the steps involved in SARS-CoV-2 infection and patho-genesis. Upon entry through the nasal route, the virus

encounters and colonizes the olfactory supporting cells present in the upper part of the nose, which turns out to be its unique host cellular target and determines its niche. The colonization is achieved by adhering to the host cells by latching the viral surface spike glycoprotein to the ACE2 receptors present in high numbers on the surface of the sustentacular cells (Flerlage *et al*. 2021). A virus, being a non-living entity, is an obligate intracellular pathogen that needs to gain access and hijack the host cellular machinery in order to replicate and, as a collateral benefit, allow them to escape immune surveillance.

The spike protein is trimeric in nature, with the three receptor binding domain (S1 subunit) forming the head, which is rather extensively glycosylated to escape recognition by antibody (Wrapp *et al*. 2020; Xia *et al*. 2020). The stalk (S2 subunit) connecting to the viral envelope is very flexible due to the presence of hinges, allowing optimal ACE2 receptor binding and thus contributes to its enhanced infectivity compared to SARS-CoV-1 (Pierri 2020; Turoňová *et al*. 2020). Post adherence, proteolysis and dissociation of the spike protein is necessary for the fusion of the viral and the host cell membrane. For this step, the virus is entirely dependent on the transmembrane serine protease (TMPRSS2), present on the host surface, to recognize and cleave the spike protein into two subunits at the poly-basic cleavage site between the S1/S2 subunits (Tang *et al*. 2020). Together with TMPRSS2, the neuropilin-1 protease found on the surface of the olfactory cells may make an alternate cleavage at the S2' hydrophobic fusion peptide site at the S2 domain

(Cantuti-Castelvetri *et al*. 2020). If the coronavirus takes the endocytic route for entry, the vesicular cathepsin-L makes these cleavages before the membrane fusion can occur (Qiu *et al*. 2006). Most of the time, the spike protein is preprocessed before new virions are even assembled. This is enabled by the furin protease, present inside the Golgi in most host cell types. Thus, a key factor in SARS-CoV-2 infection is tight recognition of the ACE2 receptor and the protease cleavage of the spike protein. This host-virus interaction also introduces tropism in coronaviruses by increasing the target space among the host cells at multiple organs, expressing the ACE2 receptor and the cell surface protease in abundance.

As the virus fulfills its sole intention, to make as many copies of it, the host-induced innate response gets triggered to reduce the pathogen burden. Sensing the presence of the viral genome, the intracellular signaling mechanism induces the interferon-stimulated genes and cytokine response, eventually leading to a shutdown of the cellular processes, cleavage of host and the viral genome by RNase-L leading to cell death (Birdwell *et al*. 2016). In addition, cytokine induced inflammatory response stops viral replication and also helps recruit phagocytes. However, the coronaviruses have evolved ways to persist by using their proteins to subvert host defense. Several of the viral non-structural proteins either bind or cleave their genome-derived triggers (dsRNA) or inhibit the host-signaling cascade (Li *et al*. 2021). This helps to antagonize the induced innate response and allows colonization in novel and normally sterile sites. The abundance of furin and other proteases in several host tissues thus promotes cellular tropism while an exuberant host inflammatory response induces tissue pathology, loss of organ function and facilitates the spread of infection. The ability to generate a high viral load in the nasopharynx and infect the host beyond the upper respiratory tract renders the coronavirus its high transmissibility.

5. Infectiousness

At the host population level, this high infectivity translates to its high infectiousness. Unlike a slowly developing chronic disease of the lower respiratory system such as tuberculosis, SARS-CoV-2 infection transits very rapidly from a susceptible state to recovery and does not persist long within an individual. Furthermore, since the nasal load achieved in the primary site of infection is enough to disseminate through viral shedding, SARS-CoV-2 also does not require optimizing its virulence with its reproduction so as to leave the body before the host dies. This also results in over-dispersion and precludes the virus from maintaining a large critical community size to persist in the population. So, now its infectiousness is determined by its 'basic reproduction number (R0)' or the number of susceptible hosts whom an infected host can pass on the disease (Endo *et al*. 2020). The ancestral Wuhan stain started with an R0 of 2.4–2.6, which increased to 3 and 4 in the European strains B.1 and B.1.1.7, respectively. It must be mentioned that the R0 number is not an intrinsic property of the virus and depends largely on host susceptibility and behavior. As a result, the recent and most successful strains of the Delta variants (B.1.617.2 and AY.1) have managed to achieve an R0 between 5 and 8 in countries with slow vaccination and low compliance to the COVID guidelines (Pueyo 2021).

6. Infectivity and immune escape of spike variants

The S1 subunit of the immunogenic spike protein has been identified as a target for many vaccines (Salvatori *et al*. 2020; Sternberg and Naujokat 2020). This subunit comprises of N-terminal domain (NTD), C-terminal domain and receptor binding domain (RBD). While the fusogenic S2 subunit contains a fusion peptide, two heptapeptide repeat sequences (HR1, HR2) and remains connected to the viral membrane with the help of a transmembrane domain (Huang *et al*. 2020). Accessibility of the spike trimer to the ACE2 receptor depends on the presence of one of the RBD in its up-state configuration, one of the three perfusion configurations that the spike protein exists (Wrapp *et al*. 2020). Cleavage at the furin site then irreversibly changes the conformation of the spike protein to a postfusion configuration, thus promoting membrane fusion. The spike protein can also undergo this transformation in the absence of binding to the ACE receptor, leading to S1 subunit shedding but no fusion (Koenig and Schmidt 2021). The S1 subunit has since accumulated multiple mutations, with NTD and the receptor binding motif being relatively more tolerant. Mutations at both of the regions significantly impacted its sensitivity to both monoclonal and vaccine stimulated neutralizing antibodies (Collier *et al*. 2021).

6.1 *D614G*

Since its identification, this mutation has become dominant in the circulating strains and is now present in every 'variant of concern'. This substitution does not directly affect the receptor binding but allows the usually disordered '630 loop' in the S1 subunit to insert into a wider

gap between the NTD and CTD created by the presence of a smaller amino acid glycine. (Zhang *et al*. 2021) Thereby stabilizing the RBD up-state configuration in the prefusion complex and preventing premature and futile S1 dissociation (Yurkovetskiy *et al*. 2020). The increased availability of the functional spike protein increases the viral infectivity but has not reported any increase in disease severity thus far.

6.2 *N501Y*

This mutation is present in the receptor binding motif of RBD. It is observed in all the highly transmissible variants like B.1.1.7, B.1.351, and P.1 (Cerutti *et al*. 2021; Faria *et al*. 2021a; Tang *et al*. 2021). N501 residue present in wild type version interacts with Y41 of ACE2 through a hydrogen bond. The substitution disrupts this interaction but allows the spike protein to have two new hydrogen bonds with ACE2 (Liu *et al*. 2021). Tighter binding of the RBD to the ACE2 receptor is the main reason for the high infectivity. In addition, the mutation also showed a significant loss of neutralization by monoclonal antibodies with reduced sensitivity to B.1.1.7 variant (Collier *et al*. 2021).

6.3 *ΔH69/ΔV70*

First identified in the Alpha variant, this deletion has been observed to co-occur with the other receptor binding motif mutations N501Y, N439K and Y453F (Meng *et al*. 2021). The mutation alone has shown to impart two-fold higher infectivity in pseudotyped lentiviral assays by increasing affinity to ACE2. The deletion has even helped rescue a binding defect observed in mutants formed *in vitro* and *in vivo* (D796H) (Kemp *et al*. 2021; Meng *et al*. 2021). Since H69-V70 is located in a disordered loop of the NTD near the recognition site of antibodies, the deletion might alter its accessibility affecting the binding of neutralizing antibodies targeting the NTD (Collier *et al*. 2021).

6.4 *E484K*

Identified first in the RBD of the B.1.351 variant. This mutation reduces the affinity of spike protein to ACE2 when present alone. E484 of the spike glycoprotein interacts with K31 of ACE2, and disruption of this interaction reduces affinity (Lan *et al*. 2020). However, when present with N501Y it increases affinity to ACE2 (Cheng *et al*. 2021). This substitution in the immunodominant epitope causes a substantial loss in binding to neutralizing antibodies induced by the two mRNA vaccines currently in use (Wang *et al*. 2021; Wibmer *et al*. 2021) and is the main reason for immune evasion to most of RDB targeting antibodies (Gobeil *et al*. 2021).

6.5 *K417N/T*

K417 lies near RBD and like E484 mutation, it reduces affinity to ACE2 when present alone but is compensated in the background of N501Y (Cheng *et al*. 2021). However, affinity to antibodies is markedly reduced due to the loss of several non-covalent interactions between the aliphatic chain of K417 with monoclonal antibodies isolated from convalescent serum.

6.6 *P681H/R*

Mutation at this position, near the S1/S2 cleavage site, was first observed in the UK variant. The mutation was predicted to affect the furin mediated cleavage but with no impact on viral entry in the Alpha variant (Lubinski *et al*. 2021). However, in the new Delta variant of concern, substitution at the same position has been observed but with an R residue (P681R). This additional arginine was found to enhance furin mediated cleavage, followed by enhanced membrane fusion.Surprisingly, the spike protein found inside and on the surface of cells infected with the B.1.617.2 (Delta) variant was predominantly in its cleaved form (Mlcochova *et al*. 2021). Cleavage of the spike prior to the release of the virion is the primary reason for it increased uptake, faster replication resulting in a thousand fold higher infectivity compared to the original Wuhan strains A/B (Li *et al*. 2021). Furthermore, like P681H, in vitro cell culture infection P681R variant showed larger syncytia formation, thus hinting at a switch in replication from cell-free infection to cell-cell fusion thus, explaining the cause of its rapid spread in the body (Saito *et al*. 2021). Thus, mutation at this position (P681H/R) is possibly associated with an increase in the fusogenic potential of the spike protein.

7. Which are the future variants of concern?

Antigenic drift leads to a gradual accumulation of mutations. We witness this each year during the flu season in the case of the influenza virus. Immune

resistance mutations at the RBD were observed in the case of the seasonal coronavirus 229E over the years (Wong *et al*. 2017) and is natural to happen for the new coronavirus as well. So, what is the future of new SARS-CoV-2 variants coming from the past? Is it possible to predict the future mutations that will escape being recognized by the current vaccine induced antibodies or will transmit rapidly in society? Bloom and colleagues have attempted to identify these future changes in the genome by using a yeast display system to produce different versions of the RBD and monitor binding to monoclonal antibodies and convalescent plasma (Greaney *et al*. 2021b). The process also helps to select mutants by growing different spike variants in the presence of neutralizing antibodies. The escape maps identified few major epitopes or hot spots in the RBD, which often acquire mutations to evade neutralization by antibodies targeting a wide surface, with the E484 residue showing the maximum impact on binding (Greaney *et al*. 2021a). The immune escape mutations can be recapitulated in the real world by studying the viral evolution as a function of convalescent plasma therapy. In an immunosuppressed individual receiving plasma therapy upon infection by the B.1 strain, a dominant viral genotype with the deletion at H69/V70 was found to emerge together with mutation (D796H) in S2 subunit (Kemp *et al*. 2021). The same mutations resulted in losing sensitivity to neutralizing monoclonal antibodies and in high infectivity when tested with pseudoviral particles. Encouragingly, treatment with non-competing monoclonal antibody combinations prevented the development of escape variants and delayed the emergence of resistance to antibodies in clinical trials (Copin *et al*. 2021). Since serum neutralization is an accepted correlate for protection, the impact caused by the individual mutations must be considered for releasing vaccine updates and when choosing a combination of monoclonal antibodies. However, we must also not forget the unchanged protection provided by the non-neutralizing (binding) antibodies and the CD8+ T cells mediated clearance of the host cells infected with SARS-CoV-2 strains (Moore *et al*. 2021; Tarke *et al*. 2021).

In conclusion, all the successful management strategies of the pandemic that we have implemented so far have been based on the knowledge of the initial SARS-CoV-2 genome sequence. This information, when collected at scale without geographical inequities, becomes pivotal in understanding how the virus is evolving and how we should change our guard in terms of its impact on accuracy in diagnosis, introducing stringent public health measures, and designing vaccination schedules. Without adequate genomic surveillance, all *in vitro* estimates of immune escape and new efforts in predicting clade emergence using machine-learning algorithms will remain theory. Until now, we have been able to sequence little more than 1% of the viral genomes from all the cases worldwide, far below the 5% recommended by the World Health Organization. It has since called for accelerated integration of genome sequencing into a regular global health practice to track the new variants and prepare for future threats from pandemics (WHO 2021a). Impactful genomic epidemiology has helped track the evolution of the latest variant of interest C.37 (Lambda), which infected 80% of the unvaccinated population in Peru with the highest mortality rate in the world. A similar effort from the South African surveillance (NGS-SA) identified a new variant C.1.2 closely related to the Lambda variant, but with distinct mutations. In addition to the well known mutations observed in the variants of concern and interest, this variant was reported to accumulate unique mutations at the N terminal domain, within the receptor binding motif and adjacent to the furin cleavage site in the spike protein. We should be concerned and monitor its immune evasion, transmissibility and the disease severity as it spreads beyond the African continents (Scheepers *et al*. 2021). Fortunately, most of the current vaccines have shown efficacy against all the SARS-CoV-2 variants. There is every reason to believe that they will protect from severe disease and save millions of lives provided we can bring them to the low-income countries. However, with more than 80% of the vaccines made available to 10% of the world's population, it will not be easy to stop generating new epicenters of variants in the developing world. Vietnam's failure in controlling the recent outbreak of the delta variant after creating the biggest success story at the beginning of the pandemic is the best example to learn from.

References

Annavajhala MK, Mohri H, Zucker JE, Sheng Z, Wang P, *et al*. 2021 A novel SARS-CoV-2 variant of concern, B.1.526 identified in New York. *medRxiv*

Birdwell LD, Zalinger ZB, Li Y, Wright PW, Elliott R, *et al*. 2016 Activation of RNase L by murine coronavirus in myeloid cells is dependent on basal oas gene expression and independent of virus-induced interferon. *J. Virol*. **90** 3160–3172

Bolze A, Cirulli ET, Luo S, White S, Wyman D, *et al*. 2021 Rapid displacement of SARS-CoV-2 variant B.1.1.7 by B.1.617.2 and P.1 in the United States. *medRxiv*

Cantuti-Castelvetri L, Ojha R, Pedro LD, Djannatian M, Franz J, *et al*. 2020 Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science*. **370** 856–860

Cerutti G, Rapp M, Guo Y, Bahna F, Bimela J, *et al*. 2021 Structural basis for accommodation of emerging B.1.351 and B.1.1.7 variants by two potent SARS-CoV-2 neutralizing antibodies. *bioRxiv*

Cheng MH, Krieger JM, Kaynak B, Arditi M and Bahar I 2021 Impact of South African 501.V2 variant on SARS-CoV-2 spike infectivity and neutralization: a structure-based computational assessment. *bioRxiv*

Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, *et al*. 2021 Convergent evolution of SARS-CoV-2 spike mutations, L452R, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. *Biorxiv*

Collier DA, De Marco A, Ferreira I, Meng B, Datir R, *et al*. 2021 SARS-CoV-2 B.1.1.7 sensitivity to mRNA vaccine-elicited, convalescent and monoclonal antibodies. *medRxiv*

Copin R, Baum A, Wloga E, Pascal KE, Giordano S, *et al*. 2021 The monoclonal antibody combination REGEN-COV protects against SARS-CoV-2 mutational escape in preclinical and human studies. *Cell*

Di H, McIntyre AA and Brinton MA 2018 New insights about the regulation of Nidovirus subgenomic mRNA synthesis. *Virology*. **517** 38–43

Elbe S and Buckland-Merrett G 2017 Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Chall*. **1** 33–46

Endo A, Abbott S, Kucharski AJ and Funk S 2020 Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open. Res*. **5** 67

Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DdS, *et al*. 2021a Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science*. **372** 815–821

Flerlage T, Boyd DF, Meliopoulos V, Thomas PG and Schultz-Cherry S 2021b Influenza virus and SARS-CoV-2: pathogenesis and host responses in the respiratory tract. *Nat. Rev. Microbiol*. **19** 425–441

Gobeil SM, Janowska K, McDowell S, Mansouri K, Parks R, *et al*. 2021 Effect of natural mutations of SARS-CoV-2 on spike structure, conformation and antigenicity. *bioRxiv*

Greaney AJ, Loes AN, Crawford KHD, Starr TN, Malone KD, *et al*. 2021a Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe*. **29** 463-476.e466

Greaney AJ, Starr TN, Gilchuk P, Zost SJ, Binshtein E, *et al*. 2021b Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe*. **29** 44-57.e49

Hamre D and Procknow JJ 1966 A new virus isolated from the human respiratory tract. *Proc. Soc. Exp. Biol. Med*. **121** 190–193

Haseltine WA 2021 A new Covid-19 variant from nigeria raises increased concerns for containment and vaccination. *https://www.forbes.com/*

Huang Y, Yang C, Xu XF, Xu W and Liu SW 2020 Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacol. Sin*. **41** 1141–1149

Kang L, He G, Sharp AK, Wang X, Brown AM, *et al*. 2021 A selective sweep in the Spike gene has driven SARS-CoV-2 human adaptation. *bioRxiv*

Kemp SA, Collier DA, Datir RP, Ferreira I, Gayed S, *et al*. 2021 SARS-CoV-2 evolution during treatment of chronic infection. *Nature*. **592** 277–282

King A 2020 An uncommon cold. *New. Sci*. **246** 32–35

Koenig PA and Schmidt FI 2021 Spike D614G—a candidate vaccine antigen against Covid-19. *N. Engl. J. Med*. **384** 2349–2351

Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, *et al*. 2020 Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*. **182** 812-827.e819

Lan J, Ge J, Yu J, Shan S, Zhou H, *et al*. 2020 Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. **581** 215–220

Li B, Deng A, Li K, Hu Y, et al. 2021 Viral infection and transmission in a large, well-traced outbreak caused by the SARS-CoV-2 Delta variant. medRxiv. *https://doi.org/10.1101/2021.07.07.21260122*

Li W, Zhang C, Sui J, Kuhn JH, Moore MJ, *et al*. 2005 Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *Embo. J*. **24** 1634–1643

Li Y, Renner DM, Comar CE, Whelan JN, Reyes HM, *et al*. 2021 SARS-CoV-2 induces double-stranded RNA-mediated innate immune responses in respiratory epithelial-derived cells and cardiomyocytes. *Proc. Natl. Acad. Sci*. **118** e2022643118

Liu H, Zhang Q, Wei P, Chen Z, Aviszus K, *et al*. 2021 The basis of a more contagious 501Y.V1 variant of SARS-CoV-2. *Cell Res*. **31** 720–722

Lubinski B, Tang T, Daniel S, Jaimes JA and Whittaker GR 2021 Functional evaluation of proteolytic activation for the SARS-CoV-2 variant B.1.1.7: role of the P681H mutation. *BioRxiv*

Meng B, Kemp SA, Papa G, Datir R, Ferreira I, et al. 2021 Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. Cell Rep. 35 109292

Mlcochova P, Kemp SA, Dhar MS, et al. 2021 SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. Nature. https://doi.org/10.1038/s41586-021-03944-y

Moore PL, Moyo-Gwete T, Hermanus T, Kgagudi P, Ayres F, et al. 2021 Neutralizing antibodies elicited by the Ad26.COV2.S COVID-19 vaccine show reduced activity against 501Y.V2 (B.1.351), despite protection against severe disease by this variant. BioRxiv

O'Toole ÁHV and Pybus OG 2021 Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2. Wellcome Open Res. 6 121

Pierri CL 2020 SARS-CoV-2 spike protein: flexibility as a new target for fighting infection. Signal Trans. Target Ther. 5 254

Pueuo T 2021 Delta variant: everything you need to know. https://unchartedterritories.tomaspueyo.com/p/delta-variant-everything-you-need

Qiu Z, Hingley ST, Simmons G, Yu C, Das Sarma J, et al. 2006 Endosomal proteolysis by cathepsins is necessary for murine coronavirus mouse hepatitis virus type 2 spike-mediated entry. J. Virol. 80 5768–5776

Rambaut A, Holmes EC, O'Toole A, Hill V, McCrone JT, et al. 2020 A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat. Microbiol. 5 1403–1407

Rambaut AL, Pybus O, Barclay W, Barrett J, Carabelli C, et al. 2020 Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. virological.org

Robertson DI 2020 First report of COVID-19 in Scotland. virological.org

Robson F, Khan KS, Le TK, Paris C, Demirbag S, et al. 2020 Coronavirus RNA proofreading: molecular basis and therapeutic targeting. Mol. Cell. 79 710–727

Saito A, Nasser H, Uriu K, Kosugi Y, Irie T, et al. 2021 SARS-CoV-2 spike P681R mutation enhances and accelerates viral fusion. bioRxiv

Salvatori G, Luberto L, Maffei M, Aurisicchio L, Roscilli G, et al. 2020 SARS-CoV-2 SPIKE PROTEIN: an optimal immunological target for vaccines. J. Transl. Med. 18 222

Scheepers C, Everatt J, Amoako DG, Tegally H, Wibmer CK, et al. 2021 Emergence and phenotypic characterization of C.1.2, a globally detected lineage that rapidly accumulated mutations of concern. medRxiv https://doi.org/10.1101/2021.08.20.21262342

Sternberg A and Naujokat C 2020 Structural features of coronavirus SARS-CoV-2 spike protein: targets for vaccination. Life Sci. 257 118056

Tang T, Bidon M, Jaimes JA, Whittaker GR and Daniel S 2020 Coronavirus membrane fusion mechanism offers a potential target for antiviral development. Antiviral Res. 178 104792

Tang JW, Toovey OTR, Harvey KN and Hui DDS 2021 Introduction of the South African SARS-CoV-2 variant 501Y.V2 into the UK. J. Infect. 82 e8–e10

Tarke A, Sidney J, Methot N, Zhang Y, Dan JM, et al. 2021 Negligible impact of SARS-CoV-2 variants on CD4+ and CD8+ T cell reactivity in COVID-19 exposed donors and vaccinees. bioRxiv

Tegally H, Wilkinson E, Lessells RJ, Giandhari J, Pillay S, et al. 2021 Sixteen novel lineages of SARS-CoV-2 in South Africa. Nat. Med. 27 440–446

Turoňová B, Sikora M, Schürmann C, Hagen WJH, Welsch S, et al. 2020 In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. Science. 370 203–208

Wall EC, Wu M, Harvey R, Kelly G, Warchal S, et al. 2021 Neutralising antibody activity against SARS-CoV-2 VOCs B.1.617.2 and B.1.351 by BNT162b2 vaccination. Lancet. 397 2331–2333

Wang P, Nair MS, Liu L, Iketani S, Luo Y, et al. 2021 Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. Nature. 593 130–135

WHO 2021a Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health

WHO 2021b Tracking SARS-CoV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/

Wibmer CK, Ayres F, Hermanus T, Madzivhandila M, Kgagudi P, et al. 2021 SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. bioRxiv

Wong AHM, Tomlinson ACA, Zhou D, Satkunarajah M, Chen K, et al. 2017 Receptor-binding loops in alphacoronavirus adaptation and evolution. Nat. Commun. 8 1735

Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, et al. 2020 Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 367 1260–1263

Xia S, Zhu Y, Liu M, Lan Q, Xu W, et al. 2020 Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. Cell Mol. Immunol. 17 765–767

Yadav PD, Sapkal GN, Ella R, Sahay RR, Nyayanit DA, et al. 2021 Neutralization of Beta and Delta variant with sera of COVID-19 recovered cases and vaccinees of inactivated COVID-19 vaccine BBV152/Covaxin. J. Travel. Med

Yurkovetskiy L, Wang X, Pascal KE, Tomkins-Tinch C, Nyalile TP, et al. 2020 Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. Cell. 183 739-751.e738

Zhang J, Cai Y, Xiao T, Lu J, Peng H, et al. 2021 Structural impact on SARS-CoV-2 spike protein by D614G substitution. Science. 372 525–530

Corresponding editor: BJ Rao