

## **Structural basis of fitness of emerging SARS-COV-2 variants and considerations for screening, testing and surveillance strategy to contain their threat.**

Sk Ramiz Islam, Debasish Prusty and Soumen Kanti Manna.

Biophysics and Structural Genomics Division, Saha Institute of Nuclear Physics (HBNI), Kolkata-700064, India.

**Abstract:** While emergence of new SARS-COV-2 variants is posing grave challenge to efforts to deal with the COVID-19 pandemic, the structural and molecular basis of their fitness remain poorly understood. We performed *in silico* analysis of structures of two most frequent SARS-COV-2 mutations, namely, N501Y and E484K, to identify plausible basis of their fitness over the original strain. The analysis suggested that the N501Y mutation is associated with strengthening of intra- as well as intermolecular H-bond in the hACE2 receptor-spike protein complex, which could result in increased affinity and, therefore, higher infectivity. While E484K mutation did not seem to directly affect the binding with hACE2 receptor, it disrupted H-bonding and salt-bridge interaction associated with binding with neutralizing antibody, which could affect chance of re-infection, disease outcome. Survey of several other mutations showing reduction in antibody-mediated neutralization also revealed that similar disruption of H-bonding or salt-bridge or Van der Waals interaction might explain their phenotype. Analysis of GESS database indicated that N501Y, E484K as well as these other mutations existed since March-April, 2020, might have evolved independently across the world and may keep accumulating, which could affect efficacy of vaccination and antibody-based therapies. Our analysis also indicated that these may spread in spite of current travel restrictions focused on few countries and evolve indigenously warranting intensification of surveillance for emerging mutations among all travellers as well as people in their dwelling zones. Meta-analysis of existing literature showed that repeat testing of travellers, contacts and others under scrutiny 7-11 days after the initial RT-PCR test may significantly help to contain the spread of emerging variants by catching false negative results. In addition, existing evidence calls for development of strain-specific tests, escalated sequencing and broadening the scope of surveillance including in hospitals and animal farms to contain the threat of emerging variants.

**Key words:** SARS-COV-2, mutation, human ACE2 receptor, neutralizing antibody, hydrogen bond, infectivity, escape variant, geographic distribution, temporal evolution, false negative RT-PCR, repeat testing, surveillance.

**Introduction:** Along with loss of life, the socioeconomic impact of COVID-19 pandemic has been devastating for a large section of world population. Several assessments indicate that it would push a vast number of people into extreme poverty, hunger and derail progress towards improving of Human Development Indices.<sup>1,2</sup> The spectrum of long-term health consequences of the SARS-COV-2 infection is quite a concern.<sup>3,4</sup> In addition, it has significantly derailed immunisation programs<sup>5</sup> increasing the danger of resurgence of preventable infectious diseases. Amid all these, the emergence of the new SARS-COV-2 strains in Britain (VUI 202012/01 or B.1.1.7) and South Africa (501Y.V2) is threatening to worsen situation.<sup>6-9</sup> The UK variant was estimated to be 71% more infectious and may increase viral load as well.<sup>6,7</sup> Emerging evidence also suggest that it may also lead to higher mortality.<sup>10</sup> It was detected in 31 countries outside UK by December, 30, 2020, itself. Almost at the same time, the variant from South Africa<sup>11</sup> spread very quickly in that country and had also been detected in travellers returning to UK. Another variant had also been also reported from Brazil in December, 2020.<sup>12</sup> While Japan reported detection of a novel variant from travellers returning from Amazonian states of Brazil in January, it was found to be different from the one circulating in Brazil.<sup>13</sup> However, understanding of the structural and molecular basis of improved fitness of these strains over parent Wuhan strain remains unclear.

As the number of cases with mutant strain kept piling up, many countries have stepped up testing and contact tracing of travellers and introduced fresh restrictions on mobility in December, 2020. However, it is quite likely that these variants actually emerged and transmitted much before that. For example, the B.1.1.7 was traced back to samples collected on September 20,<sup>7,14</sup> when none of these restrictions were in place. In fact, there has been considerable ambiguity about origin of these strains as well as the timeline of their arrival or evolution in different countries. Understanding these are very important to adopt appropriate strategy for containing the threat of these variants. In absence of adequate measures for surveillance, screening and testing, many of these strains are likely to transmit and cause havoc anywhere in the world. It is important to note that the leaky screening protocol for travellers based on body temperature and other symptoms failed us to arrest the spread of the pandemic. This kept happening in spite of emerging knowledge that many people can be asymptomatic and thus the aforementioned protocol had

significant probability of false negative. For countries except China, COVID-19 was an imported disease, which could, in principle, be controlled with rigorous testing and isolation protocols at arrival. In fact, these were quite effectively implemented in countries such as Taiwan, Singapore, and New Zealand. A similar story of failure repeated with the D614G strain, which was predicted to be more infectious than the Wuhan strain,<sup>15,16</sup> and yet could spread swiftly across the world<sup>17</sup> since testing and isolation protocols were not revamped to detect the strain. It is important to note that variants with D796H,  $\Delta$ H69/ $\Delta$ V70 mutations detected in UK<sup>18</sup> as well as E484K and E417N mutations detected in Africa is suspected to have lower susceptibility towards highly neutralizing antibodies present in convalescent plasma<sup>19</sup>. In addition to the highly infectious variants mentioned above, several other mutations have been detected in SARS-COV-2, which were found to be resistant to monoclonal antibodies<sup>20</sup> and, thus, may compromise the efficacy of vaccines and increase probability of re-infection. In fact, E484K mutation has already shown to be associated with re-infections in Brazil.<sup>21</sup> Along with travel restrictions, the current strategy to contain cross-border transmission in most countries involves requirement of RT-PCR-negative status before boarding and generic RT-PCR test on arrival and sequencing of samples turning positive. In spite of these measures and even total travel ban in many cases, imported variants are still being detected in increasing numbers around the world.<sup>22,23</sup> Thus, in this study, we analyze the structural basis of most frequent mutations in the emerging variants as well as examine the literature to find best way to reduce false negative results for travellers and analyze the spatio-temporal trend of evolution SARS-COV-2 mutants of concern to recommend expanded retrospective surveillance through strain-specific testing and escalated sequencing.

## **Method:**

***Structural implication of common mutations in emerging variants:*** The mutations found in three different variants of concern as per ECDC on its January 21, 2021 update<sup>24</sup> reported from UK (B1.1.7), South Africa (B1.351), Brazil and Japan (B1.1.28)<sup>25</sup> as well as two more strains related to infection clusters in Brazil<sup>12, 13</sup> were compared to identify most frequent mutations. Given that the most striking aspect of their phenotype has been increased infectivity and reduced susceptibility to

antibodies, we chose to analyze the effect of two most common mutations identified in the spike protein on interaction with the human ACE2 receptor and a neutralizing antibody. The high resolution crystal structure of the complex of the receptor binding domain (RBD) of the spike protein in complex with the human ACE2 receptor (hACE2) (PDB ID: 6M0J, resolution 2.6 Å)<sup>26</sup> and that with the neutralizing antibody P2B-2F6 (PDB ID: 7BWJ, resolution 2.65 Å)<sup>27</sup> were used for these analysis. Effect of individual mutations were examined analyzed by estimating the difference of binding free energies ( $\Delta\Delta G$ ) using MutaBind2 server (<https://lilab.jysw.suda.edu.cn/research/mutabind2/>).<sup>28</sup> The PDB file of the mutant complex generated by for the analysis was further used to analyze any significant deviation in the dihedral angles in the Ramachandran Plot using Ramachandran Plot Server (<https://zlab.umassmed.edu/bu/rama/>)<sup>29</sup> including glycine and proline residues. These structures were analyzed using RASMOL<sup>30</sup> (version 2.7.5.2) to examine the effect of mutations on intra- and inter-molecular interactions.

***Spatio-temporal evolution of SARS-COV-2 mutations:*** The SARS-COV-2 variant database GESS<sup>31</sup> was searched to find temporal trend of deposition and geographical distribution of single nucleotide variants (SNV) including those coding for E484 and N501Y<sup>32</sup> spike protein mutations as well as co-occurring SNVs that have been identified in emerging variants found in UK, South Africa, Brazil and Brazil-related variant outside Brazil (Japan). In addition, SNVs of spike protein RBD residues involved in H-bonding or salt bridge (K417, N440, K444, N448, N487, Q493) as well as significant Van der Waals interactions (L452, F453, A475, V483, F486, F490) with residues of neutralizing antibodies as revealed by high resolution crystal structures<sup>27, 33</sup> were also included for analysis. Moreover, along with 'escape mutations' in the spike protein RBD, others like D796H, H655Y that have been shown to be markedly resistant to antibodies<sup>18, 34-37</sup> were examined for their spatio-temporal evolution.

***Meta-analysis of rate of false negative results in RT-PCR test for SARS-COV-2:*** A search on the Pubmed with the keywords 'SARS-COV-2 + RT-PCR + false negative' was performed to find published and preprint articles. In addition, articles including preprints cited in earlier literature surveys<sup>38-40</sup> on false negative RT-PCR rates were also included. These were screened to select only original articles reporting results from at least three independent samples. These articles were

manually examined to find respective screening modalities (population-based or symptom-based or upon arrival at hospitalization), types of samples, target genes and protocols followed, percentages of false negatives in initial RT-PCR along with the basis of calling false negatives as well as interval between initial RT-PCR and reassessment to identify false negatives. Correlation between RTPCR false negative rates and the interval between initial test and reassessment through repeat RTPCR or CT scan were examined using Pearson correlation using GraphPad Prism 7. The screening of a traveller at arrival is similar in nature with population-based screening since in both cases a subject may already be infected for a while and have had a false negative result before boarding or may have been inoculated during the boarding and air travel or may still be uninfected. The concern in both cases is the ability to catch very recent and asymptomatic infections with low viral load. It is safe to assume that barring sampling and sample processing issues or RT-PCR protocol itself, most people returning false negative results may have low viral load such as those with very recent infections. The median interval between inoculation and symptom onset has been shown to be 5.1 days (95% CI, 4.5 to 5.8 days),<sup>41</sup> which was also used by Kucirka *et al*<sup>39</sup> for analysis variation of false negative results with sampling time. So, the median reassessment interval for studies reporting symptom-based testing were adjusted by an additional five days to compare them with false negatives in population-based or traveller screening associated with very recent infection. The false negative RT-PCR rates of symptom-based testing was plotted against adjusted interval together with that of population-based screening to generate the expected variation of false negative rate in the initial RT-PCR test with respect to time interval to reassessment and find the best strategy to catch these false negative results.

## **Results:**

***Identification of most frequent mutations in emerging variants:*** Table 1 summarizes changes in amino sequences identified in five different emerging variants of SARS-COV-2. While no mutation was found to be present in all of them, N501Y and E484K mutants were found in four out of five variants. So, these two

were chosen to analyze their plausible consequences on interaction with the hACE2 receptor and the neutralizing antibody P2B-2F6.

***Effect of N501Y and E484 mutations on interaction with hACE2 receptor:*** The *in silico* analysis through MutaBind2 algorithm estimated  $\Delta\Delta G$  values of 0.18 and 0.61 kcal/mole, respectively, for hACE2 complexes of E484K and N501Y mutants. Both mutations were predicted to be non-deleterious for the binding interaction. All PDB files generated for the point mutation can be found in supplementary material. The analysis of dihedral angles showed that the original crystal structure (PDB: 6M0J) had 765 residues in highly preferred conformation, 13 residues in preferred and 7 residues in questionable conformation according to the Ramachandran plot (Figure S1A). Residues with questionable conformation included three residues from the spike protein RBD and four residues from the ACE-2 receptor. Similar analysis of showed 767, 15 and 5 residues respectively, in highly preferred preferred and questionable conformation in the N501Y mutant complex (Figure S1B). The E484K mutant complex also had 767, 15 and 5 residues respectively, in highly preferred, preferred and questionable conformation (Figure S1C). In case of mutants, two residues in questionable conformation were from the ACE-2 receptor and three residues from the spike protein RBD. All residues in questionable conformation were glycine residues. Analysis showed no significant change in secondary structure of the spike protein RBD upon mutation (data not shown).

The analysis of residues involved in the interaction between hACE2 receptor and the spike protein RBD showed that N501 residue is present in the binding surface. As was shown earlier,<sup>26</sup> it was involved in interaction with the Tyr41, Lys353, Gly354 and Asp355 of hACE2 receptor. It was noted that the distance between the Y41 side chain oxygen and the side chain amide nitrogen of N501 was 3.43Å (Figure 1A), indicating a weak H-bond. Residues around this region from the RBD and the hACE2 receptor were involved in several other H-bonds that are mentioned in Table 2 and shown in Figure 1A. The Y501 residue in the N501Y mutant, on the other hand, had no H-bonding interaction with the Y41 residue (Figure 1B). But it seemed to form a strong H-bond (2.79Å) with the side chain amine of the K353 residue of the hACE2 receptor. The K353 residue was involved in a relatively weak intra-molecular H-bond (3.17Å) with one of the side chain oxygen of the D38 residue in the complex with the Wuhan strain RBD. However, this distance reduced to 2.5Å indicating a



strengthening of H-bond in the N501Y mutant. The steric bulk of the 501Y side chain was also found to cause a flip in the Q498 side chain in such a manner that the distance between the Q498 side chain amide nitrogen and the Q42 side chain amide oxygen reduced from 3.38Å for the Wuhan strain to 3.06Å in the N501Y mutant. In addition, while the distance between the backbone carbonyl oxygen of G496 residue and the N501 side chain nitrogen was 3.4Å, the 501Y side chain oxygen was placed just 2.5Å from the G496 carbonyl in the N501Y mutant.

The E484 residue is located at the surface of the spike protein RBD (Figure S2A) and was not found to be involved in any Van der Waals, H-bonding or salt-bridge interaction with the hACE2 receptor. The E484K mutation resulted in reversal of the charge of the surface exposed residue. In addition, the charged amine group of the lysine side chain was also found to be positioned on top of the aromatic ring of the F490 residue (Figure S2B). This could result in pi-cation interaction and may affect any intermolecular interaction that involves these two residues.

***Effect of N501Y and E484 mutations on interaction with neutralizing antibody:***

The MutaBind2 analysis estimated a  $\Delta\Delta G$  of 1.74 kcal/mole to be associated with the E484K mutation of the spike protein for complex formation with the P2B-2F6 neutralizing antibody. The mutation was predicted to be deleterious for the binding with P2B-2F6. The analysis of dihedral angles showed that the original crystal structure (PDB: 7BWJ) had 598 residues in highly preferred conformation, 25 residues in preferred and 7 residues were in questionable conformation according to the Ramachandran plot (Figure S3A). Residues with questionable conformation included 2, 3 and 2 residues, respectively, from the spike protein RBD, heavy and light chains of the antibody. Similar analysis showed 597, 26 and 7 residues respectively, in highly preferred, preferred and questionable conformation for the E484K mutant complex (Figure S3B). The N501Y mutant complex also had 600, 23 and 8 residues respectively, in highly preferred, preferred and questionable conformation (Figure S3C) with 2, 4 and 2 residues in questionable conformation belonging to spike protein RBD, heavy and light chains of the antibody, respectively. All residues in questionable conformation were glycine residues. Analysis showed no significant change in secondary structure of the spike protein RBD upon mutation (data not shown).



The E484 residue is involved in an H-bond interaction with the Y34 side chain OH of the heavy chain as well as a salt-bridge interaction with the R112 side chain of the light chain of the antibody (Figure 2A). Inter-atomic distances associated with intermolecular H-bond, salt-bridge and Van der Waals interactions in the vicinity of the E484 residue are mentioned in Table 3. As shown in Figure 2B, the E484K mutation flips the positively charged K484 residue away from the positively charged R112 residue of the heavy chain. Along with loss of the salt-bridge interaction, this also removes the H-bond involving the Y34 side chain of the light chain. The K484 side chain instead forms an intra-molecular H-bond (2.74Å) with the backbone carbonyl oxygen of the G485 residue. Analysis of the interface between the RBD and antibody in the vicinity of the 484 residue revealed that the mutation also weakens the H-bond (2.5Å to 3.02Å) between the 484 backbone carbonyl oxygen and side chain amide nitrogen of N33 residue of the light chain of the antibody (Figure 2B). In addition, the H-bonds involving Y27 (heavy chain) and G447 (RBD) as well as H54 (heavy chain) and N450 (RBD) were found to lengthen whereas that involving S31 (heavy chain) and Y449 (RBD) was found to shorten. The Van der Waals interaction involving the F490 residue (RBD) and V105-V106 of the heavy chain was found to be relatively unaffected.

***Spatio-temporal evolution of SARS-COV-2:*** The list of different SNVs searched in the GESS database along with their monthly as well as country-wise absolute numbers reported so far can be found in Supplementary info. As has been noted earlier, the frequency of detection of D614G mutation increased sharply from 21.74% in February to 70.53% in March, crossed 99% October, 2020 and ever since remained almost 100% across the world till date (Figure 3A). It was noted that both N501Y and E484K mutations were first reported in March, 2020 (Figure 3B). While their percentage increased relatively slowly, a spurt was noted since September, 2020. Interestingly, L18F mutation that was found in both South Africa (B1.351) and Brazil variants (P1 Manaus), could be traced back to February, 2020 and showed a sharp increase, as seen for D614G, after July, 2020. L18F mutation was significantly more abundant (> 30%) than N501Y and E484K mutations (<1%) in December, 2020. Other spike protein SNVs that have either been found in emerging variants (Table 1) or shown to be less susceptible to neutralizing antibodies were also found to be reported in the GESS database since long (Figure 3C and 3D). In

fact, except R190S, S982A all these mutations were found to be reported back in April, 2020 or earlier.

Analysis of co-occurring SNVs show that while both N501Y and E484K mutation has been reported since long, their co-occurrence (N501Y-E484K) in GESS database could only be identified since October, 2020. Nevertheless, as the sequence of events show (Figure 4C), this precedes the first reported sequencing of B.1.351 (South Africa), B.1.1.28 (outside Brazil) and P.1 Manaus (Brazil) variants as well as imposition of travel restrictions from UK, all of which happened December 2020 onwards. It was interesting to note that even according to the latest GESS database, the co-occurrences of N501Y and E484K was found only in samples from South Africa so far (see Supplementary info) while Brazilian variants harbouring them has already been sequenced. This indicated that absence in GESS should not be taken as any evidence for absence of a mutation or variant in a particular country or time. The analysis of other SNVs on RBD residues (440, 444, 448, 452, 453, 475, 483, 484, 486, 487, 490 and 493) as well as other spike protein residues (R346, H655Y) implicated in evolution of mutations with reduced susceptibility to neutralizing antibodies showed (Figure S4 and Figure 3D) that several mutations appeared throughout the pandemic and well before December, 2020.

The cumulative country-wise distribution of SNVs (Figure 5) showed that while N501Y mutation was detected in 8 countries, N501T, N501I and N501H mutations were detected in 12, 2 and 1 countries, respectively. E484K mutation was reported in 18 countries, whereas E484Q and E484A were reported in 11 and 4 countries, respectively. P681 mutation that co-occurs with N501Y in the B.1.1.7 UK variant was present in 27 countries (Figure 5S). Mutations associated with reduced susceptibility to neutralizing antibodies and escape such as K417N, H655Y, N440K, D796H, N450K were reported in 7, 23, 11, 5 and 4 countries, respectively. L452R and L452M were detected in 22 and 6 countries respectively. Other mutations associated with significant changes in ability to participate in H-bond or salt-bridge or Van der Waals interactions, such as K444R, K444N, N487D, R346I, R346S, R346K, Y453F, F490L, F490S, V483F and D138Y were detected in 3, 4, 1, 3, 1, 12, 5, 2, 7, 11 and 22 countries, respectively (Figure S6). While most of these mutations were reported from Europe, particularly, Great Britain area, South Africa, Brazil and USA, several other countries in different parts of the world showed presence of multiple mutations.

**Factors affecting false negative results in RTPCR test for SARS-COV-2:** A search on the Pubmed with the keywords 'SARS-COV-2 + RT-PCR + false negative' returned 223 results including published and preprint articles. Along with original articles found in them, published articles and preprints cited in review articles in Pubmed were also included. These articles were screened to identify articles reporting analysis of at least three samples clearly mentioning subject selection criteria for initial RT-PCR, basis for calling false negative and percentage of false negative results. This yielded 24 original articles including four on population-based screening,<sup>42-45</sup> ten on symptomatic test<sup>46-55</sup> and ten on tests performed at hospitalization<sup>56-65</sup> as summarized in Tables 4, 5 and 6 respectively. The percentage of false negative results for the initial RT-PCR test that were reported in these studies ranged between 2-100%. The study reporting 100% false negative result was performed on day one at hospitalization. The weighted average of false negative rates for the initial RT-PCR analysis in population screening (122399 samples), symptomatic testing (1692 samples) and hospital-based testing (2742 samples) were estimated to be 8.41%, 9.97% and 17.45%, respectively. The overall false negative rates combining all of them were estimated to be 8.58%. Among population-based sampling studies, it was noted that as the time gap between initial RT-PCR and the repeat RT-PCR or CT scan to confirm results increased from 1 day to 4 days to 6.1 days on average, the detection rates of false negatives increased linearly (Figure 6A,  $R^2 = 0.878$ ) from 2% to 4% to 9.3% (Table 4).<sup>42-44</sup> Data from the population-based screening study with sero-positivity-based false negative calling<sup>45</sup> could not be used in analysis since a significant number of people may not become sero-positive. On the other hand, false negative rate for symptom-based RT-PCR tests decreased linearly (Figure 6B,  $R^2 = 0.674$ ) with re-assessment interval wherever they were mentioned (see Table 5).<sup>46, 48, 51, 54, 55</sup> While three studies on hospital-based testing mentioning the interval to reassessment<sup>57, 58, 63</sup> showed no such linearity between reassessment interval and false negative rate (Figure 6C, Table 6), repeat RT-PCR testing of nasal swabs after a median interval of 7.5 days reported detection of highest fraction of initial false negative RT-PCR results.<sup>58</sup> Plot of RTPCR false negative detection on reassessment against adjusted interval between initial RT-PCR and reassessment combining population-based screening<sup>42-44</sup> and symptom-based testing studies<sup>46, 48, 51, 54, 55</sup> showed a peak between 7-8 days. It should be noted that although desirable, the interval between symptom onset and initial RT-

PCR could not be accounted for in the adjusted time interval since none of the above studies mentioned that clearly. Data from hospital-based tests could also not be combined since interval between infection or symptom onset and hospitalization is highly variable and context-dependent. The plot indicated that repeat test after 15 days of initial RT-PCR could reduce the identification of false negative results below 10%.

## **Discussion:**

Mutation is an integral part of the evolution of a pathogenic virus. The most challenging consequences of a mutation include increase in infectivity, decrease in susceptibility to neutralizing antibodies due antigenic drift as well as decrease in sensitivity towards detection by existing methods. The SARS-COV-2 enters the host cell by high affinity attachment through hACE2 receptor with a dissociation constant of 4.7 nM.<sup>27</sup> Our analysis showed that the N501Y mutation may further enhance this affinity through a stronger inter-molecular H-bonding network. The 501, 498 and 496 residues are located in a relatively flexible region of the RBD between an alpha helical turn and a beta-sheet. Thus the enhanced intramolecular H-bonding involving the 496 and Y501 may also provide conformational stability to allow such stronger inter-molecular interactions leading to stronger affinity and increased infectivity associated with emerging variants from UK, South Africa and Brazil harbouring N501Y mutation. However, the fact that this mutation has been there since March 2020, indicates either inadequate sequencing of samples associated with higher infectivity or mechanistic contribution of other mutations that needs to be investigated. Along with increase in hACE2 affinity and viral load, contribution of other mutations in B.1.1.7 variant in affecting the pathobiology should be examined to explain the suspected increase in mortality. In addition, recent results indicate that N501Y may itself affect interaction with some antibodies.<sup>66</sup>

The spike protein RBD also binds to neutralizing antibodies such as P2B-2F6,<sup>27</sup> C002, C121, C135 and C144 generated by immune system.<sup>33</sup> However, the neutralizing antibodies seem to approach the RBD at a different angle (compared to the hACE2) that does not involve interaction with 501 residue. Earlier study showed that the spike RBD and P2B-2F6 interacts with a dissociation constant of 5.14 nM<sup>27</sup>

making the neutralizing antibody competitive against hACE2, which is essential to prevent infection. The interaction involves H-bond, salt-bridge as well as Van der Waals interactions involving several residues at the RBD-antibody interface. E484K mutation disrupts the salt bridge and H-bonding interaction with R112 of the heavy chain as well as H-bond with Y34 and N33 of the light chain. The Coulombic repulsion between R112 and K484 also weakens several other intermolecular H-bonds. Complex with C002 shows E484 side chain interacting with R96 (heavy chain) and H35 (light chain), both with positively charged side chains.<sup>27</sup> So, similar to that observed in case P2B-2F6, K484 side chain would feel the Coulombic repulsion and disrupt these H-bonds. C121-RBD complex involves H-bond between side chain -OH groups of Y33 and S52 (heavy chain) with same E484 side chain oxygen,<sup>27</sup> which has to serve as a donor in one and acceptor in the other H-bond. The C144-RBD complex also shows the E484 side chain oxygen in similar simultaneous donor acceptor role in H-bond with S53 and S56 residue of the heavy chain.<sup>27</sup> E484K mutation would disrupt such interactions as well given that the lysine side-chain is two carbons longer and the side chain amine of the lysine is likely to be protonated and hence serve poorly as an acceptor. Thus, E484K mutation would compromise neutralizing activity of several antibodies and 'escape' as has been documented.<sup>35</sup> On the other hand, C135 neutralizing antibody interacts with N440 side chain of RBD via H-bond network involving D53 and R55 of the heavy chain.<sup>27</sup> Their interaction also engages R346 side chain of RBD in H-bond with backbone carbonyls of Y98 (heavy chain) and Y91 (light chain) as well as a pi-cation interaction with the W32 (light chain) aromatic ring.<sup>27</sup> This would explain the observed reduction in susceptibility associated with N440K and R346S mutations.<sup>35</sup>

<sup>36</sup> In general, mutations causing significant changes in side chain polarity, charge as well as length or steric bulk of RBD residues that may participate in intermolecular H-bond (E484, K417, K444, N448, N487, Q493, Q498, N440, R346), salt bridge (E484, K417, K444, R346) as well as significant Van der Waals interactions (L452, Y453, A475, V483, F486, F490) are expected to compromise the neutralizing activity. Thus, it was not surprising that earlier studies have found mutations like E484K, K417N as well as L452R, A475V and V483A to result in reduced vulnerability to neutralizing antibodies presumably due to disruption of H-bond/salt-bridge and Van der Waals interactions, respectively. In fact, an earlier study on in vitro evolution of RBD as well as non-RBD spike protein variants showed that all escape mutants were associated

with either change in side chain charge and/or size and hydrophobicity.<sup>36</sup> D796H and H655Y mutations in the spike protein, which has also been associated to compromised susceptibility to neutralizing antibodies,<sup>18, 36</sup> also involve reversal or removal of side chain charges and are indicative of these residues being involved in H-bond/salt-bridge interaction with antibody. It would be interesting to examine if the L18F mutation found in the B1.1.7 variant contributes to suggested increase in mortality due to disruption of Van der Waals interaction involving the L18 residue. Such mutations significantly affecting spike-antibody interaction could compromise the ability of a person previously infected with the Wuhan strain to neutralize the mutant strain. In case the new variant also harbours a mutation strengthening binding with the competing hACE2 receptor (e.g., N501Y), the neutralizing antibodies could lose decisively leading to re-infection as has been documented for the P1 Manaus strain from Brazil harbouring E484K along with N501Y mutation.<sup>67</sup> Co-occurrences and accumulation of these mutations could elevate reinfection risk as well as pose challenge to vaccination and antibody-based therapies.

It is interesting to note that while E484K and other mutations mentioned above could be traced back to March-April, 2020. Also, E484K, K417N and other mutations associated with reduced antibody susceptibility and escape from neutralization are already present in many countries outside Brazil, South Africa or UK. However, re-infections related to them have been documented only recently and still remains modest in number. One of the reasons might have been due to inadequate sequencing and/or detection of the first infection, which may have been asymptomatic. The GISAID database is currently overrepresented by sequences from developing countries. It does not represent the actual geographical distribution of SAR-COV-2 genotypes since many developing countries including India, Brazil, Russia and South Africa has very high number of cases per sequence deposited from those countries. Thus, it is possible that there is a hidden variant load in many of these countries. This is presumably a result of the fact that sequencing is both time-consuming and resource-intensive. However, given that mutations of concern have been reported since March-April, 2020 and are already present in several countries suggest that the possibility of their independent emergence and evolution into strains of concern in different parts of the world is quite real. The difference in mutation profile of two Brazil-related variants identified is also indicative independent



branches of evolution existing side-by-side. Novel strains with N501Y has now been detected in Ohio,<sup>68</sup> USA, whereas, P681H detected in Nigeria,  $\Delta$ 69-70HV detected in Connecticut and Wisconsin, USA<sup>70</sup> and L452R detected in Southern California<sup>71</sup> independently. It's interesting to note that while sample containing the UK B1.1.7 variant were first collected in September 20, 2020, the first set of UK-focused travel restrictions by EU members were imposed two months after that. Several other countries followed that and escalated sequencing of samples from COVID-positive travellers from UK. So, at least for two months, there was possibility of people with the variant travelling across country and continents.

It should be noted that even when the whole world was under lockdown and recent COVID-negative status was mandatory to fly; several people were being detected as COVID-positive upon arrival. It's not surprising that with the requirement of recent COVID-negative results to fly even now, several travellers from UK and Europe are being detected with the SARS-COV-2 including the mutant strain. For example, India has now detected 150 cases of the UK variant.<sup>23</sup> One plausible reason could be that these people picked the infection somewhere on their way to boarding or on board from a fellow traveller. The other possibility is that pre-boarding RT-PCR tests yielded false negative results allowing a COVID-positive person onboard. It should be remembered that this person as well as inoculated fellow travellers, may return a false negative result even at arrival and then spread the strain in the country of destination. For example, the reported B1.1.7 tally in India increased from 58 in January 5<sup>72</sup> to 150 on January 23<sup>23</sup> in spite of ongoing travel restriction and mandatory testing on arrival. Thus, there is every possibility that persons carrying novel strains, such as B.1.1.7 did deliver disseminate them to new destinations during September 20, 2020 and December 20, 2020, when EU countries re-imposed travel restrictions from UK.<sup>73</sup> Also, given that travel restrictions have been UK-centric to start with, they may not have prevented spreading of other strains of concern as well as the B1.1.7 strain from a third country.

So, along with escalated sequencing, measures to reduce false negatives results and to track persons with false negative RT-PCR are very important. In case there are no issues with sampling, sample processing issues or RT-PCR primer and protocol, false negatives may arise due to inadequate viral load. A passenger may have low viral load at arrival either due to being infected just during air travel, which



would not generally 1-2 days, or because of low dose exposure and better innate immune response. Most of such patients are expected to be asymptomatic as well. A pooled analysis estimated median probability of false negative result to be lowest (20%) at day 8 post-infection, while it is was 66% on day 21 and 100% on day 1.<sup>39</sup> Our analysis assuming median incubation period of 5 days indicated that repeat testing after 7-8 days of initial RT-PCR would result in detection of highest number of false negatives in the initial RT-PCR. A recent meta-analysis revised the median incubation period to 7.7 days,<sup>74</sup> which would extend the repeat test window to 9.7-10.7 days. So, a repeat test at 7-11 days after the initial RTPCR at arrival along with the current requirement of RT-PCR-negative status immediately before boarding would significantly help to reduce false negative results and reduce transmission. However, as our analysis indicate, reassessment beyond 15 days may significantly compromise the ability to identify initial false negative results.

Given that mutants of concern seem to be quite well-spread around the world and may be evolving further, it would be prudent to warrant repeat assessment for all international travellers. It would be best to keep them in quarantine for two weeks at least till their reassessment results are available. However, it should be noted that it may not be universally acceptable to quarantine all travellers for that long. So, in case the RT-PCR on arrival is negative, which upon repeat test turns out to be false negative, and the traveller is allowed to leave the airport, the variant can spread, particularly if the subject dwells in a densely populated area. Thus, irrespective of initial RT-PCR results, all travellers should be counselled to come under digital surveillance to identify their dwell zone till the initial negative result is confirmed through re-assessment. This would help in monitoring any unusual spike in disease dynamics in these areas and take appropriate actions in case the person is eventually found to harbour a variant of concern through sequencing. It should, however, be noted that some mutations may increase the likelihood for false negative results in RT-PCR.<sup>75</sup> In fact, the  $\Delta H69/\Delta V70$  mutation was found to affect the result of diagnostic RTPCR tests with S target gene.<sup>10</sup> Thus, the sensitivity of the strain of interest should be thoroughly monitored and multiplexed PCR protocols should be adopted.<sup>76</sup> Development of strain-specific primers<sup>77</sup> and RT-PCR protocols should be taken up expeditiously so that surveillance for strains interest

can be improved even without mandatory sequencing, which is time-consuming, expensive and requires specific expertise.

In view of the fact that large number asymptomatic travellers with false negative tests may have already escaped detection under the current protocol of test at arrival, travellers from countries reporting emerging variants dating back to, at least, September, 2020 should be catalogued and their dwell zones should be mapped. These areas should be monitored and the temporal evolution of COVID case load, severity and mortality maps should be compared with dwell zone maps to examine any correlation, which should warrant escalated testing and sequencing effort.

Mutation and antigenic drift is an integral part of the evolution of a pathogenic virus. While it may be interesting that the emerging variant B1.1.7 has acquired fourteen mutations over the Wuhan strain, we should have seen it coming. It should not be any surprise for a virus that apparently could accumulate several mutations, including an insertion at the functionally crucial S1/S2 cleavage site, in its spike protein alone over its bat variant to infect humans.<sup>78</sup> In fact, its ability to swiftly mutate and transmit back and forth between human and animals is manifest from the SARS-COV-2 mutations including N501T, F486L and Y453F in the spike protein RBD detected in connection with infections in mink farms.<sup>79</sup> The RBD mutations like Y453F was also found in cats whereas N501T was found in ferrets and mice were found to have K417N, Q493H/K, Q498H as well as N501Y mutations.<sup>80</sup> A recent study showed presence of N501T in a human variant along with Q493K that went undetected and has been estimated to have evolved in Italy in early August, 2020.<sup>81</sup> Our analysis showed N501T to be first reported in March, 2020 and now present in 12 countries including Italy with counts increasing steadily. Our analysis also showed that the N501Y as well as E484K, K417N and other mutations that have been shown to be associated with reduced susceptibility to antibodies have been detected as early as in March-April, 2020. These mutations are already present in several countries across the world besides UK, Brazil and South Africa. Double SNVs such as N501Y/K484K or N501Y/K417N found in the Brazilian and South African variants were detected as early as in October, 2020. These are indicative of the fact that these mutations have been accumulating in SARS-COV-2 and circulating around for quite some time. They only raised alarm and got identified since their cluster stood out against the background of a decreasing number of cases, probably, associated

with an earlier and relatively less infectious strain. It is interesting to note that out of five emerging strains of concern; only the one reported from South Africa had D614G mutation in spite of the fact that D614G has been a super-dominant mutation across the world since March, 2020. It may be indicative of independent and isolated evolution of other four variants which may precede that of the South African variant. The other possibility is a very rapid evolution one of the very few non-D614G strain bypassing transmissional advantage of the D614G strain.<sup>15,16</sup> This would be possible if it evolves within the same host. It has been shown that the virus can evolve within immunocompromised patients<sup>82</sup>. Deletion in the Y145 region has been found in immunocompromised patients and it was speculated that the B.1.1.7 may have initially evolved in such patients.<sup>83</sup> This indicates to possibility of emergence of distinct variants without necessary linkage with dominant strains around. It is important to note immunocompromised patients also show prolonged shedding.<sup>82, 84</sup> Apart from patients under treatment for conditions like cancer, organ transplant, autoimmune diseases, SARS-COV-2 patients also receive immunosuppressive agents like corticosteroids. Thus, hospitals with such patients are potentials hot zones for SARS-COV-2 evolution and should be under rigorous surveillance and escalated sequencing. In fact, in January 2021, a cluster of 35 new infections has been detected in a hospital in Germany<sup>85</sup> potentially belonging to a new strain that remains to be sequenced. Given the agility of SARS-COV-2 to evolve to and fro between animals and humans, animals farms should also be under surveillance to detect and contain emergence of any new variant at the earliest.

In conjunction with existing knowledge about other viruses like influenza,<sup>86, 87</sup> we should also expect further mutations and antigenic drift during SARS-COV-2 evolution across the world. Mutations with simple increase in infectivity without any bearing on immune response are likely to increase overall number of casualties and may also increase in mortality among old and vulnerable. However, 'escape mutations', like E484K,<sup>35-37</sup> leading to reduced efficacy of neutralizing antibodies are more likely to increase disease severity, duration and poor outcome even among healthier patients. In addition, reduced efficacy of antibodies would increase risk of re-infection along with possibility of antibody-dependent enhancement in previously infected or vaccinated individuals. Thus, any significant spike in hospitalization and disease severity or duration among young and healthier individuals should be

examined as a possible sign of emergence of a variant of concern. Since younger people are more active and many of them may still be asymptomatic or mildly symptomatic, they can spread the variant fast, especially, in areas with high population density. So, surveillance mechanism should allow examination of demographic distribution of evolving scenario. For any suspected case, similar to that in case of travellers, repeat RT-PCR test should be performed 7-11 days after the initial test.

While an earlier version of this work with similar recommendations was submitted to the Lancet (January 4, 2021) for consideration as a correspondence article, WHO expert group met on January 12 to give a series of recommendations on increasing surveillance, testing, sequencing and calling for evidence-based strategy to mitigate challenges faced by emerging variants.<sup>88</sup> Subsequently, a correspondence article on action plan for pan-EU defence against emerging variants was published. Elaborate recommendations included increased surveillance, sequencing and specific mention of repeat test 7-10 days after travel along with RT-PCR negative status within 24 hours prior to travel and 10 days strict quarantine.<sup>89</sup> Our analysis showed the threat of emerging variants are more widespread than currently assumed and may not be contained with prospective travel restrictions and testing alone. Some of these mutants may already have crossed borders in absence of appropriate testing or sequencing protocols and many more may independently emerge in different parts of the world, particularly, in hospitals, animal farms and densely populated areas. It is noted with concern that a recent analysis revealed that N501Y, E484K and K417N mutants show significant resistance to antibodies elicited by mRNA vaccines that are currently being used for immunization.<sup>66</sup> Analysis and specific suggestions made in this article should help to take a more calibrated approach towards surveillance, socio-economic activities, vaccination and antibody-based therapies to deal with the pandemic in real time.

### **Summary:**

This study revealed the structural basis of increased infectivity and reduced antibody sensitivity associated with emerging variants. It showed that several mutations associated with such effects emerged very early during the pandemic and may have

been evolving independently. It analyzed the available data to arrive at the following specific recommendations to deal with challenges of emerging variants. (1) At least one repeat RTPCR testing of travellers or contacts or any suspected individual 7-11 days after the initial RT-PCR test at arrival or referral, (2) digital surveillance on dwell zone of travellers for COVID-19 dynamics in case mandatory two weeks quarantine cannot be imposed, (3) retrospective cataloguing of international travellers from countries reporting emerging variants and analysis of COVID-19 dynamics in their dwell zones, (4) escalation in sequencing and development of variant-specific and -sensitive diagnostic method, (5) surveillance of hospital and animal farms, (6) monitoring of any unusual spike in disease burden, severity and outcome among healthy, young individuals, (7) monitoring of any spike in re-infection or adverse health problems in vaccinated individuals.

**Supplementary Material:** Supplementary figures and legends can be found in supplementary material.

**Conflict of Interest:** The author has no conflict of interest to declare.

**Acknowledgement:** Authors are supported by funding from the Department of Atomic Energy, Government of India.

#### **References:**

1. <https://www.un.org/development/desa/en/news/sustainable/sustainable-development-goals-report-2020.html>
2. Khetrapal S, Bhatia R. Impact of COVID-19 pandemic on health system & Sustainable Development Goal 3. Indian J Med Res 2020;151:395-9
3. [https://www.who.int/docs/default-source/coronaviruse/risk-comms-updates/update-36-long-term-symptoms.pdf?sfvrsn=5d3789a6\\_2](https://www.who.int/docs/default-source/coronaviruse/risk-comms-updates/update-36-long-term-symptoms.pdf?sfvrsn=5d3789a6_2)
4. Del Rio C, Collins LF, Malani P. Long-term Health Consequences of COVID-19. JAMA. 2020;324(17):1723-1724
5. Santoli JM, Lendley MC, deSilva MB, Kharbanda EO, Daley MF, Galloway L, Gee J, Glover M, Herring B, Kang Y, Lucas P, Noblit C, Tropper J, Vogt T, Weintraub E. Effects of the COVID-19 Pandemic on Routine Pediatric

- Vaccine Ordering and Administration — United States, 2020.  
<https://www.cdc.gov/mmwr/volumes/69/wr/mm6919e2.htm>
6. New and Emerging Respiratory Virus Threats Advisory Group. NERVTAG meeting on SARS-CoV-2 variant under investigation VUI-202012/01. 18 Dec 2020. <https://khub.net/documents/135939561/338928724/SARS-CoV-2+variant+under+investigation%2C+meeting+minutes.pdf/962e866b-161f-2fd5-1030-32b6ab467896?t=1608470511452>
  7. Mahase E. Covid-19: What have we learnt about the new variant in the UK? *BMJ* 2020; 371
  8. <https://www.who.int/csr/don/21-december-2020-sars-cov2-variant-united-kingdom/en/>
  9. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/955239/NERVTAG\\_paper\\_on\\_variant\\_of\\_concern\\_VO\\_C\\_B.1.1.7.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/955239/NERVTAG_paper_on_variant_of_concern_VO_C_B.1.1.7.pdf)
  10. <https://www.who.int/csr/don/31-december-2020-sars-cov2-variants/en/>
  11. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, Doolabh D, Pillay S, San EJ, Msomi N, Mlisana K, von Gottberg A, Walaza S, Allam A, Ismail A, Mohale T, Glass AJ, Engelbrecht S, Van Zyl G, Preiser W, Petruccione F, Sigal S, Hardie D, Marais G, Hsiao M, Korsman S, Davies MA, Tyers L, Mudau I, York D, Maslo C, Goedhals D, Abrahams S, Laguda-Akingba O, Alisoltani-Dehkordi A, Godzik A, Wibmer CK, Sewell BT, Lourenço J, Alcantara LCJ, Pond SLK, Weaver S, Martin D, Lessells RJ, Bhiman JN, Williamson C, de Oliveira T. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. <https://www.medrxiv.org/content/10.1101/2020.12.21.20248640v1>
  12. Genomic characterization of a novel SARS-CoV-2 lineage from Rio de Janeiro, Brazil. Voloch CM, da Silva F Jr R, de Almeida LGP, Cardoso CC, Brustolini OJ, Gerber AL, de C Guimarães AP, Mariani D, da Costa RM, Ferreira Jr OC, Covid19-UFRJ Workgroup, LNCC-Workgroup, Cavalcanti AC, Frauches TS, de Mello CMB, Galliez RM, Faffe DS, M P P Castiñeiras TMPP, Tanuri A, R de Vasconcelos ATR. <https://www.medrxiv.org/content/10.1101/2020.12.23.20248598v1>

13. Phylogenetic relationship of SARS-CoV-2 sequences from Amazonas with emerging Brazilian variants harboring mutations E484K and N501Y in the Spike protein. Naveca F, Nascimento V, Souza V, Corado A, Nascimento F, Silva G, Costa A, Duarte D, Pessoa K, Gonçalves L, Brandão MJ, Jesus M, Fernandes C, Pinto R, de Silva M, Mattos T, Wallau GL, Siqueira MM, Resende PC, Delatorre E, Gräf T, Bello G. <https://virological.org/t/phylogenetic-relationship-of-sars-cov-2-sequences-from-amazonas-with-emerging-brazilian-variants-harboring-mutations-e484k-and-n501y-in-the-spike-protein/585>
14. Rambaut A, Loman N, Pybus O, Barclay W, Barrett J, Carabelli A, Connor T, Peacock T, Robertson DL, Volz E. COVID-19 Genomics Consortium UK (CoG-UK). Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. 18 Dec 2020. <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563>
15. Laha S, Chakraborty J, Das S, Manna SK, Biswas S, Chatterjee R. Characterizations of SARS-CoV-2 mutational profile, spike protein stability and viral transmission. *Infect Genet Evol.* 2020 Nov;85:104445.
16. Zhang L, Jackson CB, Mou H, Ojha A, Peng H, Quinlan BD, Rangarajan ES, Pan A, Vanderheiden A, Suthar MS, Li W, IZard T, Rader C, Farzan M, Choe H. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat Commun.* 2020 Nov 26;11(1):6013.
17. <https://www.lanl.gov/updates/sars-cov-2-variant.php>
18. Kemp SA, Collier DA, Datir R, Ferreira I, Gayed S, Jahun A, Hosmillo M, Rees- Spear C, Mlcochova P, Lumb IU, Roberts DJ, Chandra A, Temperton N, Sharrocks K, Blane E, Briggs J, van Gils MJ, Smith K, Bradley JR, Smith C, Doffinger R, Ceron-Gutierrez L, Barcenas-Morales G, Pollock DD, Goldstein RA, Smielewska A, Skittrall JP, Gouliouris T, Goodfellow IG, Gkrania-Klotsas E, Illingworth C, McCoy LE, Gupta RK. Neutralising antibodies in Spike mediated SARS-CoV-2 adaptation. medRxiv [Preprint]. 2020 Dec 29:2020.12.05.20241927. doi: 10.1101/2020.12.05.20241927.
19. Andreano E, Piccini G, Licastro D, Casalino L, Johnson NV, Paciello I, Monego SD, Pantano E, Manganaro N, Manenti A, Manna R, Casa E, Hyseni



- I, Benincasa L, Montomoli E, Amaro RE, McLellan JS, Rappuoli R. SARS-CoV-2 escape *in vitro* from a highly neutralizing COVID-19 convalescent plasma. bioRxiv [Preprint]. 2020 Dec 28:2020.12.28.424451. doi: 10.1101/2020.12.28.424451. PMID: 33398278; PMCID: PMC7781313.
20. Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, Zhao C, Zhang Q, Liu H, Nie L, Qin H, Wang M, Lu Q, Li X, Sun Q, Liu J, Zhang L, Li X, Huang W, Wang Y. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. Cell. 2020 Sep 3;182(5):1284-1294.e9. doi: 10.1016/j.cell.2020.07.012. Epub 2020 Jul 17. PMID: 32730807; PMCID: PMC7366990.
21. Resende PC, Bezerra JF, Vasconcelos RHT, Arantes I, Appolinario L, Mendonça AC, Paixão AC, Rodrigues AC, Silva T, Rocha AS, Pauvolid-Corrêa A, Motta FC, Teixeira DLF, de Oliveira Carneiro TF, Neto FPF, Herbster ID, Leite AB, Riediger IN, Debur MC, Naveca FG, Almeida W, Livorati M, Bello G, Siqueira MM. Spike E484K mutation in the first SARS-CoV-2 reinfection case confirmed in Brazil, 2020. Virological.org (2020). <https://virological.org/t/spike-e484k-mutation-in-the-first-sars-cov-2-reinfection-case-confirmed-in-brazil-2020/584>
22. <https://www.bbc.com/news/uk-55786409>
23. <https://www.thehindu.com/news/national/150-people-infected-with-uk-variant-of-covid-19-in-india-says-government/article33642745.ece>
24. <https://www.ecdc.europa.eu/en/publications-data/covid-19-infographic-mutations-current-variants-concern>
25. Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings. Faria NR, Claro IM, Candido D, Franco LAM, Andrade PA, Coletti TM, Silva CAM, Sales FC, Manuli ER, Aguiar RS, Gaburo N, Camilo CC, Fraiji NA, Crispim MAE, Carvalho MPSS, Rambaut A, Loman N, Pybus OG, Sabino EC, on behalf of CADDE Genomic Network. <https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586>
26. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X. Structure of the SARS-CoV-2 spike receptor-binding domain

- bound to the ACE2 receptor. *Nature*. 2020 May;581(7807):215-220. doi: 10.1038/s41586-020-2180-5. Epub 2020 Mar 30. PMID: 32225176.
27. Ju B, Zhang Q, Ge J, Wang R, Sun J, Ge X, Yu J, Shan S, Zhou B, Song S, Tang X, Yu J, Lan J, Yuan J, Wang H, Zhao J, Zhang S, Wang Y, Shi X, Liu L, Zhao J, Wang X, Zhang Z, Zhang L. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature*. 2020 Aug;584(7819):115-119.
28. Zhang N, Chen Y, Lu H, Zhao F, Alvarez RV, Goncarenco A, Panchenko AR, Li M. MutaBind2: Predicting the Impacts of Single and Multiple Mutations on Protein-Protein Interactions. *iScience*. 2020 Mar 27;23(3):100939.
29. Anderson RJ, Weng Z, Campbell RK, Jiang X. Main-Chain Conformational Tendencies of Amino Acids. *Proteins*. 2005 Sep 1;60(4):679-89.
30. Bernstein HJ. Recent changes to RasMol, recombining the variants. *Trends Biochem Sci*. 2000 Sep;25(9):453-5.
31. Fang S, Li K, Shen J, Liu S, Liu J, Yang L, Hu CD, Wan J. GESS: a database of global evaluation of SARS-CoV-2/hCoV-19 sequences. *Nucleic Acids Res*. 2021 Jan 8;49(D1):D706-D714
32. Leung K, Shum MH, Leung GM, Lam TT, Wu JT. Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020. *Euro Surveill*. 2021 Jan;26(1):2002106. doi: 10.2807/1560-7917.ES.2020.26.1.2002106. PMID: 33413740; PMCID: PMC7791602.
33. Barnes CO, Jette CA, Abernathy ME, Dam KA, Esswein SR, Gristick HB, Malyutin AG, Sharaf NG, Huey-Tubman KE, Lee YE, Robbiani DF, Nussenzweig MC, West AP Jr, Bjorkman PJ. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature*. 2020 Dec;588(7839):682-687.
34. Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, Zhao C, Zhang Q, Liu H, Nie L, Qin H, Wang M, Lu Q, Li X, Sun Q, Liu J, Zhang L, Li X, Huang W, Wang Y. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell*. 2020 Sep 3;182(5):1284-1294.e9.

35. Weisblum Y, Schmidt F, Zhang F, DaSilva J, Poston D, Lorenzi JC, Muecksch F, Rutkowska M, Hoffmann HH, Michailidis E, Gaebler C, Agudelo M, Cho A, Wang Z, Gazumyan A, Cipolla M, Luchsinger L, Hillyer CD, Caskey M, Robbiani DF, Rice CM, Nussenzweig MC, Hatziioannou T, Bieniasz PD. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants
36. Baum A, Fulton BO, Wloga E, Copin R, Pascal KE, Russo V, Giordano S, Lanza K, Negron N, Ni M, Wei Y, Atwal GS, Murphy AJ, Stahl N, Yancopoulos GD, Kyratsous CA. Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science*. 2020 Aug 21;369(6506):1014-1018
37. Liu Z, VanBlargan LA, Rothlauf PW, Bloyet LM, Chen RE, Stumpf S, Zhao H, Errico JM, Theel ES, Ellebedy AH, Fremont DH, Diamond MS, Whelan SPJ. Landscape analysis of escape variants identifies SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *bioRxiv* [Preprint]. 2020 Nov 8:2020.11.06.372037. doi: 10.1101/2020.11.06.372037. PMID: 33442690; PMCID: PMC7805447.
38. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, Del Campo R, Ciapponi A, Sued O, Martinez-García L, Rutjes AW, Low N, Bossuyt PM, Perez-Molina JA, Zamora J. False-negative results of initial RT-PCR assays for COVID-19: A systematic review. *PLoS One*. 2020 Dec 10;15(12):e0242958. doi: 10.1371/journal.pone.0242958. PMID: 33301459; PMCID: PMC7728293.
39. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. *Ann Intern Med*. 2020 May 13 : M20-1495.
40. Woloshin S, Patel N and Kesselheim AS. False Negative Tests for SARS-CoV-2 Infection — Challenges and Implications. *N Engl J Med* 2020; 383:e38
41. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J. The Incubation Period of Coronavirus Disease 2019

- (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med.* 2020 May 5;172(9):577-582
42. Kanji JN, Zelyas N, MacDonald C, Pabbaraju K, Khan MN, Prasad A, Hu J, Diggle M, Berenger BM, Tipples G. False negative rate of COVID-19 PCR testing: a discordant testing analysis. *Virol J.* 2021 Jan 9;18(1):13. doi: 10.1186/s12985-021-01489-0. PMID: 33422083; PMCID: PMC7794619.
43. Long DR, Gombar S, Hogan CA, Greninger AL, O'Reilly-Shah V, Bryson-Cahn C, Stevens B, Rustagi A, Jerome KR, Kong CS, Zehnder J, Shah NH, Weiss NS, Pinsky BA, Sunshine J. Occurrence and Timing of Subsequent SARS-CoV-2 RT-PCR Positivity Among Initially Negative Patients. *Clin Infect Dis.* 2020 Jun 7 ciae722. doi: 10.1093/cid/cia 722. Epub ahead of print. PMID: 32506118; PMCID: PMC7314163.
44. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology.* 2020 Aug;296(2): E32-E40. doi: 10.1148/radiol.2020200642. Epub 2020 Feb 26. PMID: 32101510; PMCID: PMC7233399.
45. Baron RC, Risch L, Weber M, Thiel S, Grossmann K, Wohlwend N, Lung T, Hillmann D, Ritzler M, Bigler S, Egli K, Ferrara F, Bodmer T, Imperiali M, Heer S, Renz H, Flatz L, Kohler P, Vernazza P, Kahlert CR, Paprotny M, Risch M. Frequency of serological non-responders and false-negative RT-PCR results in SARS-CoV-2 testing: a population-based study. *Clin Chem Lab Med.* 2020 Aug 31;58(12):2131-2140. doi: 10.1515/cclm-2020-0978. PMID: 32866113.
46. He JL, Luo L, Luo ZD, Lyu JX, Ng MY, Shen XP, Wen Z. Diagnostic performance between CT and initial real-time RT-PCR for clinically suspected 2019 coronavirus disease (COVID-19) patients outside Wuhan, China. *Respir Med.* 2020 Jul; 168:105980. doi: 10.1016/j.rmed.2020.105980. Epub 2020 Apr 21. PMID: 32364959; PMCID: PMC7172864.
47. Duan X, Guo X, Qiang J. A retrospective study of the initial 25 COVID-19 patients in Luoyang, China. *Jpn J Radiol.* 2020 Jul;38(7):683-690. doi: 10.1007/s11604-020-00988-4. Epub 2020 May 26. PMID: 32458125; PMCID: PMC7249986.
48. An J, Liao X, Xiao T, Qian S, Yuan J, Ye H, Qi F, Shen C, Wang L, Liu Y, Cheng X, Li N, Cai Q, Wang F, Chen J, Li G, Cai Q, Liu Y, Wang Y, Zhang F,

- Fu Y, He Q, Tan X, Liu L, Zhang Z. Clinical characteristics of recovered COVID-19 patients with re-detectable positive RNA test. *Ann Transl Med.* 2020 Sep;8(17):1084. doi: 10.21037/atm-20-5602. PMID: 33145303; PMCID: PMC7575971.
49. Long C, Xu H, Shen Q, Zhang X, Fan B, Wang C, Zeng B, Li Z, Li X, Li H. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? *Eur J Radiol.* 2020 May; 126:108961. doi: 10.1016/j.ejrad.2020.108961. Epub 2020 Mar 25. PMID:32229322; PMCID: PMC7102545.
50. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for Typical Coronavirus Disease 2019 (COVID-19) Pneumonia: Relationship to Negative RT-PCR Testing. *Radiology.* 2020 Aug;296(2): E41-E45. doi: 10.1148/radiol.2020200343. Epub 2020 Feb 12. PMID: 32049601; PMCID: PMC7233363.
51. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, Ji W. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology.* 2020 Aug;296(2):E115-E117. doi:10.1148/radiol.2020200432. Epub 2020 Feb 19. PMID: 32073353; PMCID: PMC7233365.
52. Bernheim A, Mei X, Huang M, Yang Y, Fayad ZA, Zhang N, Diao K, Lin B, Zhu X, Li K, Li S, Shan H, Jacobi A, Chung M. Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection. *Radiology.* 2020 Jun;295(3):200463. doi: 10.1148/radiol.2020200463. Epub 2020 Feb 20. PMID:32077789; PMCID: PMC7233369.
53. Chen D, Jiang X, Hong Y, Wen Z, Wei S, Peng G, Wei X. Can Chest CT Features Distinguish Patients with Negative from Those with Positive Initial RT-PCR Results for Coronavirus Disease (COVID-19)? *AJR Am J Roentgenol.* 2021 Jan;216(1):66-70. doi: 10.2214/AJR.20.23012. Epub 2020 May 5. PMID: 32368928.
54. Besutti G, Giorgi Rossi P, Iotti V, Spaggiari L, Bonacini R, Nitrosi A, Ottone M, Bonelli E, Fasano T, Canovi S, Colla R, Massari M, Lattuada IM, Trabucco L, Pattacini P; Reggio Emilia COVID-19 Working Group. Accuracy of CT in a cohort of symptomatic patients with suspected COVID-19 pneumonia during the outbreak peak in Italy. *Eur Radiol.* 2020 Dec;30(12):6818-6827. doi:10.1007/s00330-020-07050-x. Epub 2020 Jul 14. PMID: 32666316; PMCID: PMC7358325.

55. Chen ZH, Li YJ, Wang XJ, Ye YF, Wu BL, Zhang Y, Xuan WL, Bao JF, Deng XY. Chest CT of COVID-19 in patients with a negative first RT-PCR test: Comparison with patients with a positive first RT-PCR test. *Medicine (Baltimore)*. 2020 Jun 26;99(26): e20837. doi: 10.1097/MD.00000000000020837. PMID: 32590775; PMCID: PMC7328934.
56. Yuan J, Kou S, Liang Y, Zeng J, Pan Y, Liu L. Polymerase Chain Reaction Assays Reverted to Positive in 25 Discharged Patients With COVID-19. *Clin Infect Dis*. 2020 Nov 19;71(16):2230-2232. doi: 10.1093/cid/ciaa398. PMID: 32266381; PMCID: PMC7184423.
57. Li C, Luo F, Xie L, Gao Y, Zhang N, Wu B. Chest CT study of fifteen COVID-19 patients with positive RT-PCR retest results after discharge. *Quant Imaging Med Surg*. 2020 Jun;10(6):1318-1324. doi: 10.21037/qims-20-530. PMID: 32550140; PMCID: PMC7276372.
58. Xie C, Lu J, Wu D, Zhang L, Zhao H, Rao B, Yang Z. False negative rate of COVID-19 is eliminated by using nasal swab test. *Travel Med Infect Dis*. 2020 Sep-Oct; 37:101668. doi: 10.1016/j.tmaid.2020.101668. Epub 2020 Apr 11. PMID: 32283215; PMCID: PMC7151360.
59. Li Y, Yao L, Li J, Chen L, Song Y, Cai Z, Yang C. Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID-19. *J Med Virol*. 2020 Jul;92(7):903-908. doi: 10.1002/jmv.25786. Epub 2020 Apr 5. PMID: 32219885; PMCID: PMC7228231.
60. Fechner C, Strobel K, Treumann T, Sonderegger B, Azzola A, Fornaro J, Schrading S, Roos JE. COVID-19 and the role of imaging: early experiences in Central Switzerland. *Swiss Med Wkly*. 2020 Jun 22;150: w20304. doi: 10.4414/smw.2020.20304. PMID: 32580213.
61. Lan FY, Filler R, Mathew S, Buley J, Iliaki E, Bruno-Murtha LA, Osgood R, Christophi CA, Fernandez-Montero A, Kales SN. COVID-19 symptoms predictive of healthcare workers' SARS-CoV-2 PCR results. *PLoS One*. 2020 Jun 26;15(6):e0235460. doi: 10.1371/journal.pone.0235460. PMID: 32589687; PMCID: PMC7319316.
62. Chen HJ, Qiu J, Wu B, Huang T, Gao Y, Wang ZP, Chen Y, Chen F. Early chest CT features of patients with 2019 novel coronavirus (COVID-19) pneumonia: relationship to diagnosis and prognosis. *Eur Radiol*. 2020



- Nov;30(11):6178-6185.doi: 10.1007/s00330-020-06978-4. Epub 2020 Jun 9.  
PMID: 32518987; PMCID:PMC7280678
- 63.Çinkooğlu A, Hepdurgun C, Bayraktaroğlu S, Ceylan N, Savaş R. CT imaging features of COVID-19 pneumonia: initial experience from Turkey. *Diagn Interv Radiol.* 2020 Jul;26(4):308-314. doi: 10.5152/dir.2020.20307. PMID: 32558645;PMCID: PMC7360080.
- 64.Dai H, Zhang X, Xia J, Zhang T, Shang Y, Huang R, Liu R, Wang D, Li M, Wu J,Xu Q, Li Y. High-resolution Chest CT Features and Clinical Characteristics of Patients Infected with COVID-19 in Jiangsu, China. *Int J Infect Dis.* 2020 Jun; 95:106-112. doi: 10.1016/j.ijid.2020.04.003. Epub 2020 Apr 6. PMID:32272262; PMCID: PMC7136866.
- 65.Lan L, Xu D, Ye G, Xia C, Wang S, Li Y, Xu H. Positive RT-PCR Test Results in Patients Recovered From COVID-19. *JAMA.* 2020 Apr 21;323(15):1502-1503. doi:10.1001/jama.2020.2783. PMID: 32105304; PMCID: PMC7047852.
- 66.Wang Z, Schmidt F, Weisblum Y, Muecksch F, Barnes CO, Finkin S, Schaefer-Babajew D, Cipolla M, Gaebler C, Lieberman JA, Yang Z, Abernathy MA, Huey-Tubman KE, Hurley A, Turroja M, West KA, Gordon K, Millard KG, Ramos V, Da Silva J, Xu J, Colbert RA, Patel R, Dizon J, Unson-O'Brien C, Shimeliovich I, Gazumyan A, Caskey M, Bjorkman PJ, Casellas R, Hatzioannou T, Bieniasz PD, Nussenzweig MC. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants
- 67.SARS-CoV-2 reinfection by the new Variant of Concern (VOC) P.1 in Amazonas, Brazil. Naveca F, da C, Nascimento V, Souza V, Corado A, Nascimento F, Costa A, Duarte D, Silva G, Mejía M, Pessoa K, Gonçalves L, Brandão MJ, Jesus M, Pinto R, Silva M, Mattos T, Abdalla L, Santos JH, Costa-Filho R, Gabriel Luz Wallau JL, Siqueira MM, Delatorre E, Gräf T, Bello G, Resende PC. <https://virological.org/t/sars-cov-2-reinfection-by-the-new-variant-of-concern-voc-p-1-in-amazonas-brazil/596>
- 68.Tu H, Avenarius MR, Kubatko L, Hunt M, Pan X, Ru P, Garee J, Thomas K, Mohler P, Pancholi P, Jones D. Distinct Patterns of Emergence of SARS-CoV-2 Spike Variants including N501Y in Clinical Samples in Columbus Ohio



69. Oluniyi PE, et al. Detection of SARS-CoV-2 P681H Spike Protein Variant in Nigeria
70. Moreno G, Braun K, Larsen BB, Alpert T, Worobey M, Grubaugh N, Friedrich T, O'Connor D, Fauver J, and Brito A. Detection of non-B.1.1.7 Spike  $\Delta$ 69/70 sequences (B.1.375) in the United States.
71. Zhang W, Davis BD, Chen SS, Sincuir Martinez JM, Plummer JT, Vail E. Emergence of a novel SARS-CoV-2 strain in Southern California, USA. <https://www.medrxiv.org/content/10.1101/2021.01.18.21249786v1.full.pdf>
72. <https://www.hindustantimes.com/india-news/nine-more-people-test-positive-for-the-new-uk-variant-of-sars-cov-2-total-reaches-38/story-aZwO2FKK5QMveKHaBdkp4J.html>
73. <https://www.thinkglobalhealth.org/article/updated-timeline-coronavirus>
74. Qin J, You C, Lin Q, Hu T, Yu S, Zhou XH. Estimation of incubation period distribution of COVID-19 using disease onset forward time: A novel cross-sectional and forward follow-up study. *Sci Adv.* 2020 Aug 14;6(33):eabc1202.
75. Artesi M, Bontems S, Göbbels P, Franckh M, Maes P, Boreux R, Meex C, Melin P, Hayette MP, Bours V, Durkin K. A Recurrent Mutation at Position 26340 of SARS-CoV-2 Is Associated with Failure of the E Gene Quantitative Reverse Transcription-PCR Utilized in a Commercial Dual-Target Diagnostic Assay. *J Clin Microbiol.* 2020 Sep 22;58(10):e01598-20
76. Peñarrubia L, Ruiz M, Porco R, Rao SN, Juanola-Falgarona M, Manissero D, López-Fontanals M, Pareja J. Multiple assays in a real-time RT-PCR SARS-CoV-2 panel can mitigate the risk of loss of sensitivity by new genomic variants during the COVID-19 outbreak. *Int J Infect Dis.* 2020 Aug;97:225-229.
77. Lopez-Rincon A, Perez-Romero CA, Tonda A, Mendoza-Maldonado L, laassen E, Garssen J, Kraneveld AD. Design of Specific Primer Set for Detection of B.1.1.7SARS-CoV-2 Variant using Deep Learning.

78. Zhou H, Chen X, Hu T, Li J, Song H, Liu Y, Wang P, Liu D, Yang J, Holmes EC, Hughes AC, Bi Y, Shi W. A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein.
79. van Dorp L, Tan CSC, Lam SD, Richard D, Owen C, Berchtold D, Orango C, Balloux F. Recurrent mutations in SARS-CoV-2 genomes isolated from mink point to rapid host-adaptation. <https://www.biorxiv.org/content/10.1101/2020.11.16.384743v1>
80. Gary RF. Mutations arising in SARS-CoV-2 spike on sustained human-to-human transmission and human-to-animal passage. <https://virological.org/t/mutations-arising-in-sars-cov-2-spike-on-sustained-human-to-human-transmission-and-human-to-animal-passage/578>
81. Fiorentini S, Messali S, Zani A, Caccuri F, Giovanetti M, Ciccozzi M, Caruso A. First detection of SARS-CoV-2 spike protein N501 mutation in Italy in August, 2020. *Lancet Infect Dis.* 2021 Jan 12;S1473-3099(21)00007-4.
82. Choi B, Choudhary MC, Regan J, Sparks JA, Padera RF, Qiu X, Solomon IH, Kuo HH, Boucau J, Bowman K, Adhikari UD, Winkler ML, Mueller AA, Hsu TY, Desjardins M, Baden LR, Chan BT, Walker BD, Lichterfeld M, Brigl M, Kwon DS, Kanjilal S, Richardson ET, Jonsson AH, Alter G, Barczak AK, Hanage WP, Yu XG, Gaiha GD, Seaman MS, Cernadas M, Li JZ. Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. *N Engl J Med.* 2020 Dec 3;383(23):2291-2293.
83. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/947048/Technical\\_Briefing\\_VOC\\_SH\\_NJL2\\_SH2.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/947048/Technical_Briefing_VOC_SH_NJL2_SH2.pdf)
84. Avanzato VA, Matson MJ, Seifert SN, Pryce R, Williamson BN, Anzick SL, Barbian K, Judson SD, Fischer ER, Martens C, Bowden TA, de Wit E, Riedo FX, Munster VJ. Case Study: Prolonged Infectious SARS-CoV-2 Shedding from an Asymptomatic Immunocompromised Individual with Cancer. *Cell.* 2020 Dec 23;183(7):1901-1912.e9.

85. <https://www.cnbc.com/2021/01/19/new-covid-strain-germany-discovers-covid-variant-in-bavaria.html>
86. Shao W, Li X, Goraya MU, Wang S, Chen JL. Evolution of influenza a virus by mutation and re-assortment. *International journal of molecular sciences*. 2017 Aug;18(8):1650.
87. Boni MF. Vaccination and antigenic drift in influenza. *Vaccine*. 2008 Jul 18; 26(Suppl 3): C8–14
88. [https://www.who.int/news/item/15-01-2021-statement-on-the-sixth-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-coronavirus-disease-\(covid-19\)-pandemic](https://www.who.int/news/item/15-01-2021-statement-on-the-sixth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic)
89. Priesemann V, Balling R, Brinkmann MM, Ciesek S, Cypionka T, Eckerle I, Giordano G, Hanson C, Hel Z, Hotulainen P, Klimek P, Nassehi A, Peichl A, Perc M, Petelos E, Prainsack B, Szczurek E.. An action plan for pan-European defence against new SARS-CoV-2 variants. *Lancet*. 2021 Jan 21;S0140-6736(21)00150-1

## **Figure Legends:**

**Figure 1.** Intra-and intermolecular H-bonding in the (A) Wuhan strain and (B) N501Y mutant of SARS-COV-2 spike protein receptor binding domain in complex with human ACE2 receptor. Blue ribbon indicates human ACE2 protein chain while cyan ribbon in panel A and red in panel B indicate spike protein RBDs of Wuhan strain and N501Y mutants, respectively. Red, pink, orange and green dotted lines indicate new, lost, shortened and lengthened H-bond, respectively. Inter-atomic distances indicated here are mentioned in Table 2.

**Figure 2.** Intra-and intermolecular H-bonding involving the (A) Wuhan strain and (B) E484K mutant of SARS-COV-2 spike protein receptor binding domain in complex with P2B-2F6 neutralizing antibody. Pink, red and green dotted lined indicate lost, new and weakened H-bonds, respectively. Blue, yellow and cyan ribbons indicate spike protein RBD, heavy chain (H) and light chains (L) of the P2B-2F6 neutralizing antibody in the Wuhan strain complex. Heavy and light chains in the E484K complex are represented by green and red ribbons. Spacefill models showing positions of (C) N501 and (D) Y501 residues of spike protein RBD in respective complexes with P2B-2F6 neutralizing antibody. The colour scheme for RBD and light chain, heavy chain of the antibody is same as that in A and B.

**Figure 3.** Monthly variation in percentage of SNVs representing (A) D614G and (B) N501Y, E484K and L18F mutations in spike protein in GESS database. Monthly variation in number of sequences with SNVs representing (C) R190S, R246I, A475V, V483A, A570D, D716I, S982A, T1027I, D1118H and (D) P26S, D80A, D138Y, D215G, D417N, H655Y, P681H, A701V in GESS database.

**Figure 4.** Monthly variation in number of sequences with co-occurring SNV pairs involving (A) E484K or N501 or D614 mutations and (B) L18F and D614G mutation as reported in GESS database. (C) Sequence of events related to emergence mutations of concern.

**Figure 5.** Country-wise relative abundance of N501Y, N501T, N501Y, N501H and E484K, E484Q and E484A mutations as reported in GESS database.

**Figure 6.** Plot of rate of detection of false negative results in initial RT-PCR against time interval to re-assessment based on repeat RT-PCR or CT scan in (A) population based SARS-COV-2 screening studies, (B) symptom-based testing and (C) hospitalization-based testing. (D) Variation of rate of detection of false negative results in initial RT-PCR against adjusted time interval to reassessment.

### **Supplementary Figure Legends:**

**Figure S1.** Analysis of dihedral angle of amino acid residues and their position in the Ramachandran plot for (A) Wuhan strain (PDB: 6M0J), (B) N501Y mutant and (C) E484K mutant of SARS-COV-2 spike protein receptor binding domain bound to the human ACE2 receptor. PDB files for mutant complexes (see supplementary material) generated via MutaBind2 server. Green, orange and red dots indicates residues in highly preferred, preferred and questionable conformation. Only glycine residues were found to be in questionable conformation.

**Figure S2.** Spacefill models showing surface exposed (A) E484 residue of the spike RBD domain of the Wuhan strain (shown in cyan) and (B) K484 residue of the E484K mutant (in red) of SARS-COV-2 along with F490 residue in complex with the human ACE2 receptor (in blue). The F490 residue below the

**Figure S3.** Analysis of dihedral angle of amino acid residues and their position in the Ramachandran plot for (A) Wuhan strain (PDB: 7BWJ), (B) N501Y mutant and (C) E484K mutant of SARS-COV-2 spike protein receptor binding domain in complex with P2B-2F6 neutralizing antibody. PDB files for mutant complexes (see supplementary material) generated via MutaBind2 server. Green, orange and red dots indicates residues in highly preferred, preferred and questionable conformation. Only glycine residues were found to be in questionable conformation.

**Figure S4.** Monthly variations in SNVs detected in (A) 346, 417, 444, (B) 448, 450, 452, 453, 475, 483, 484, (C) 486, 487, 490 and 493 positions of spike protein RBD shown to be involved in interaction with neutralizing antibodies.

**Figure S5.** Figure S5. Country-wise relative abundance of P681H, K417N, H655Y, N440K, D796H, N450K, L452R and L452M mutations as reported in GESS database.

**Figure S6.** Country-wise relative abundance of K444R, K444N, N487D, R346I, R346S, R346K, Y453F, F490L, F490S, V483F, D138Y mutations as reported in GESS database.

Figure 1

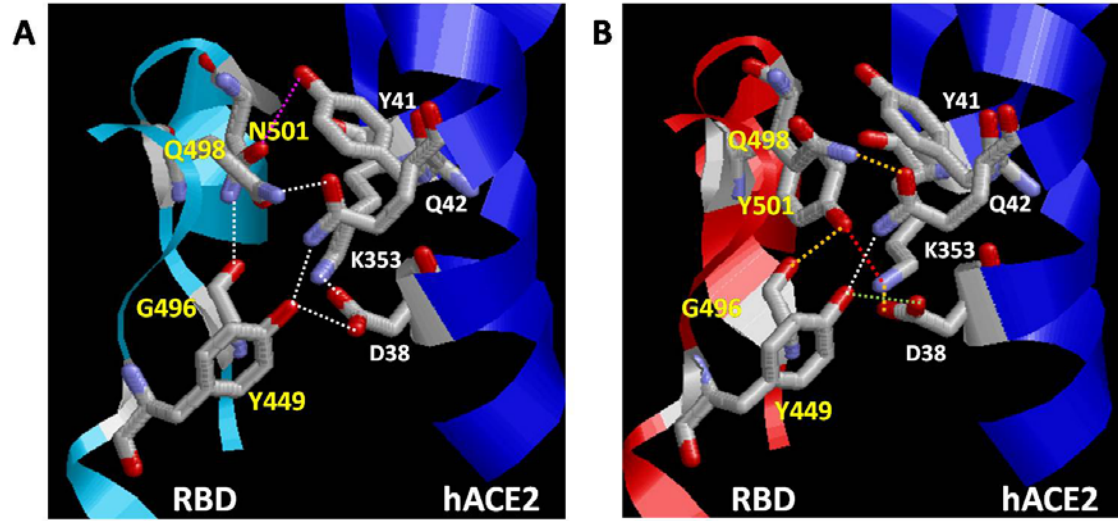




Figure 2

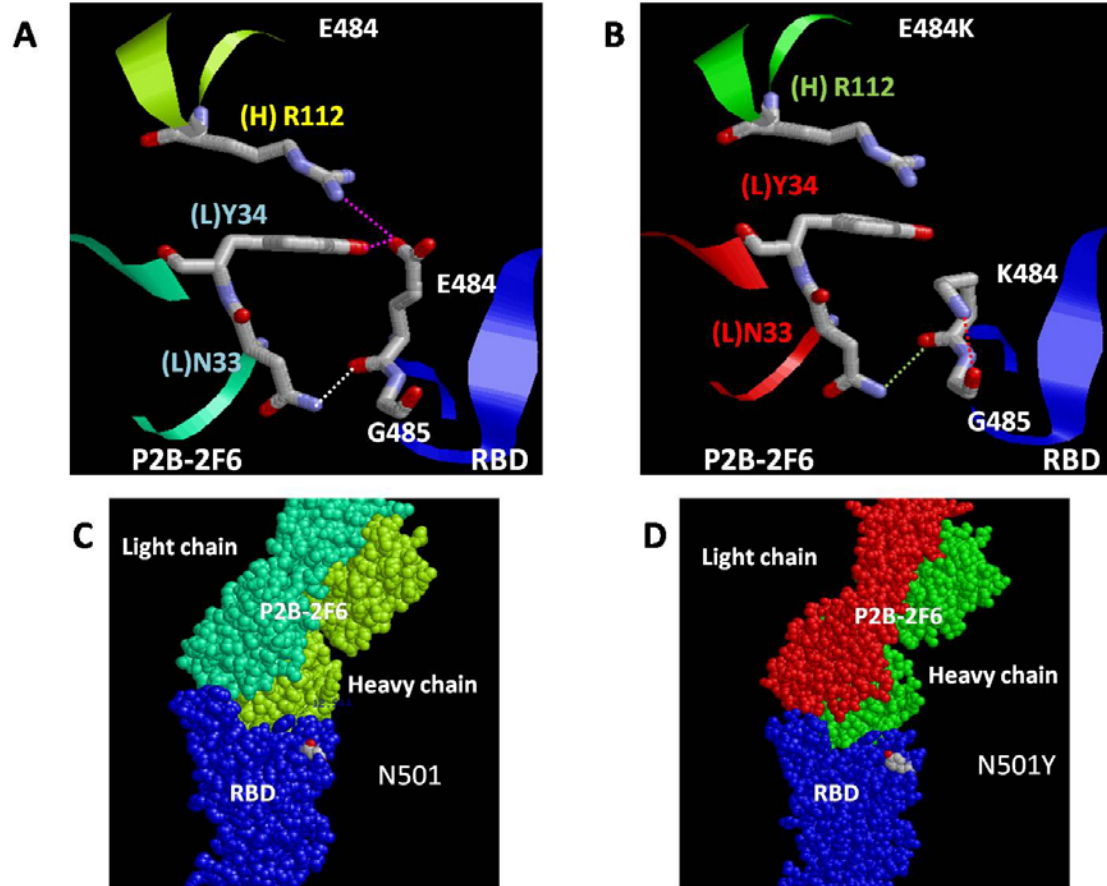


Figure 3

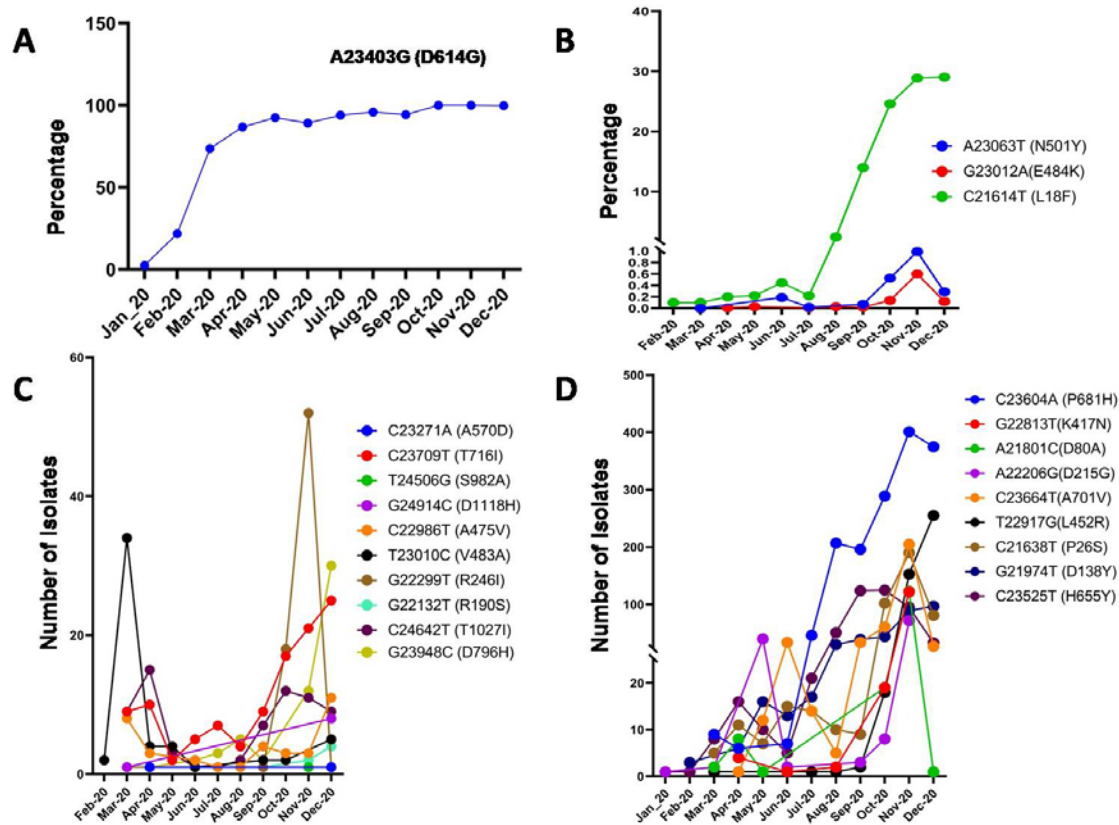


Figure 4

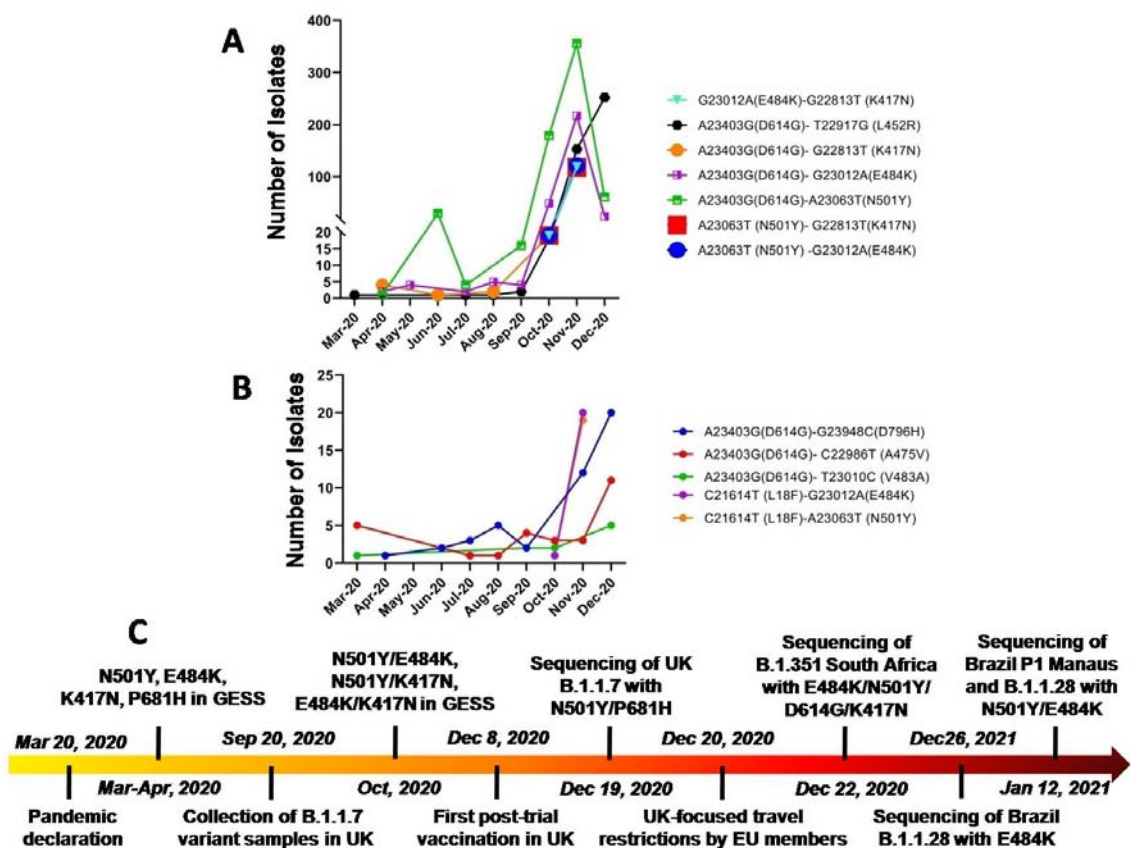


Figure 5

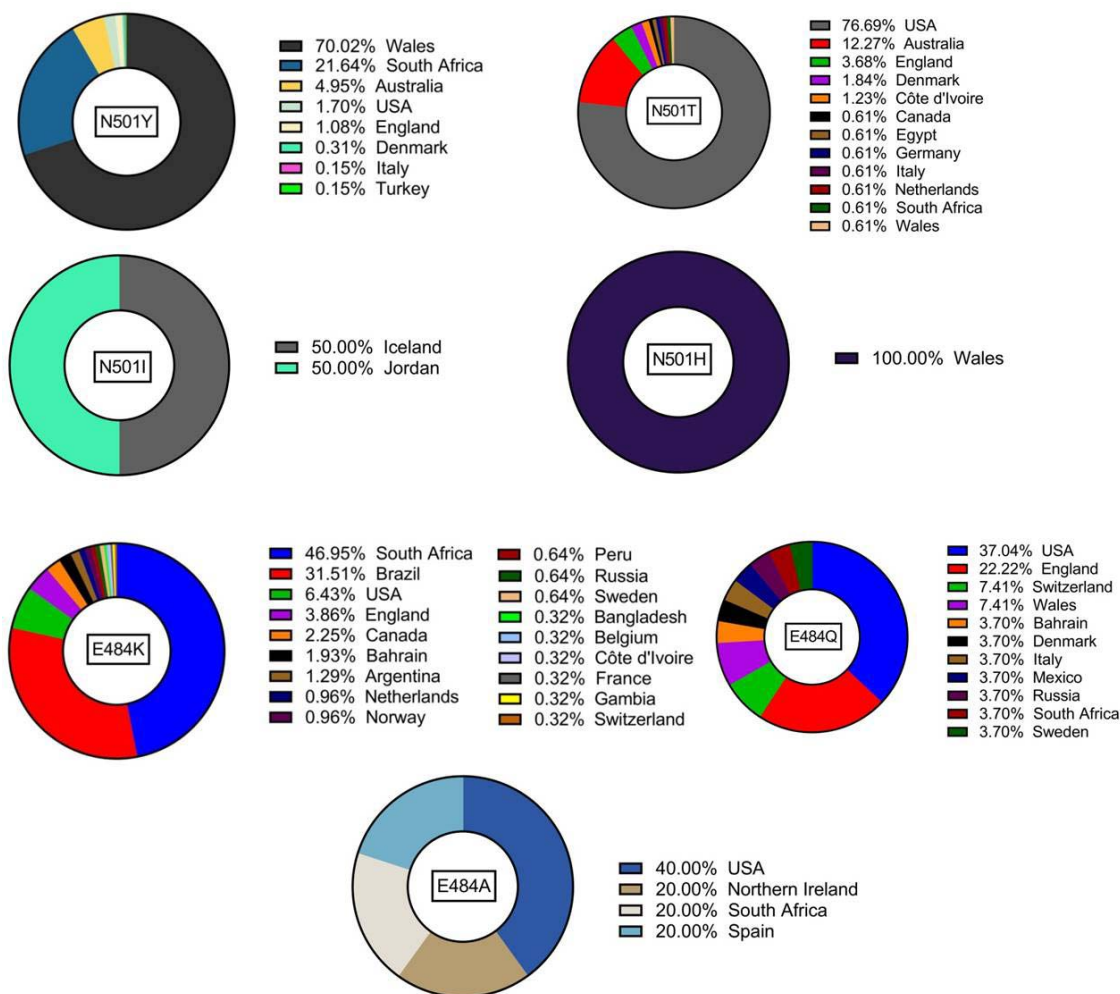


Figure 6

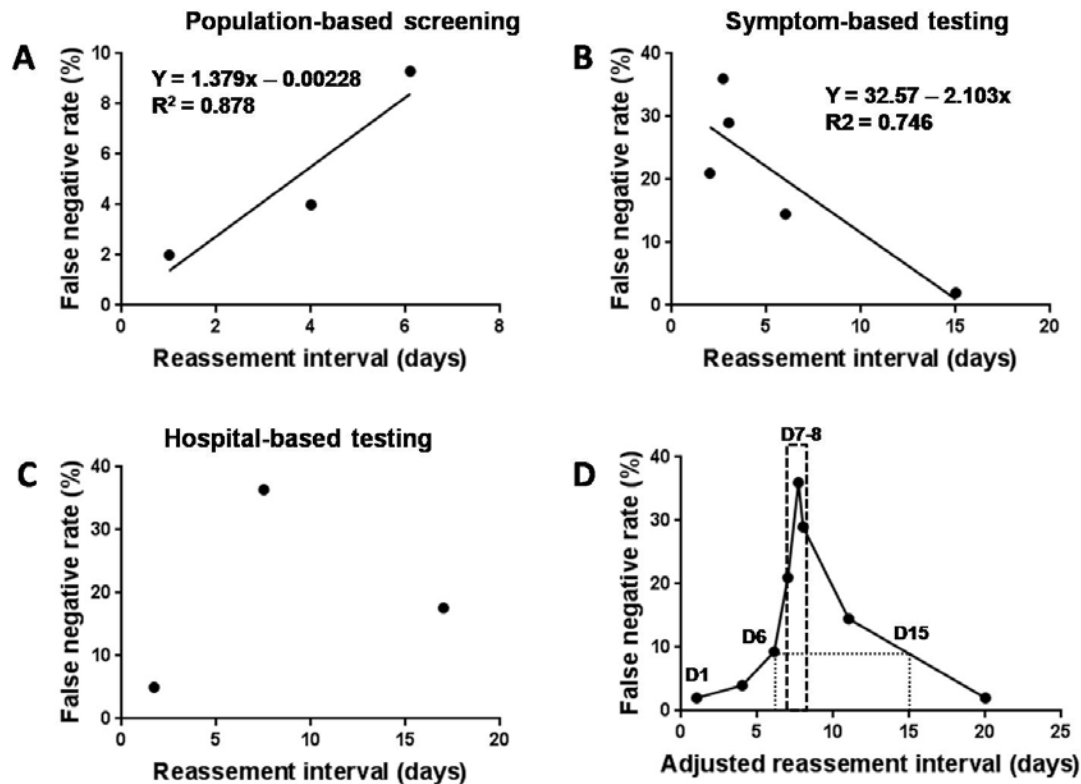




Figure S1

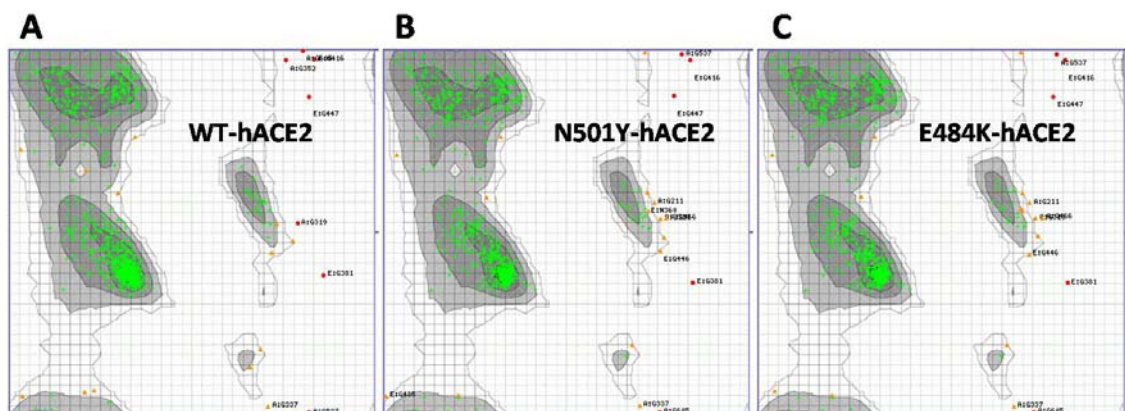


Figure S2

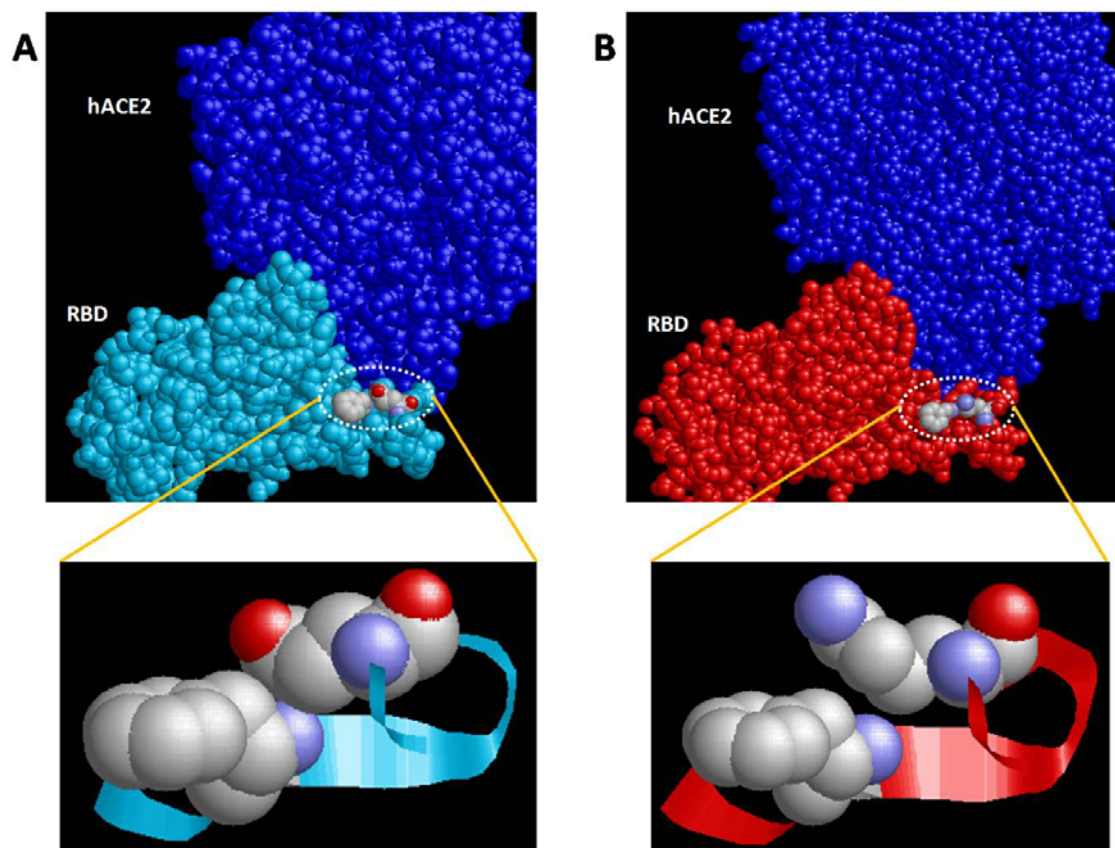


Figure S3

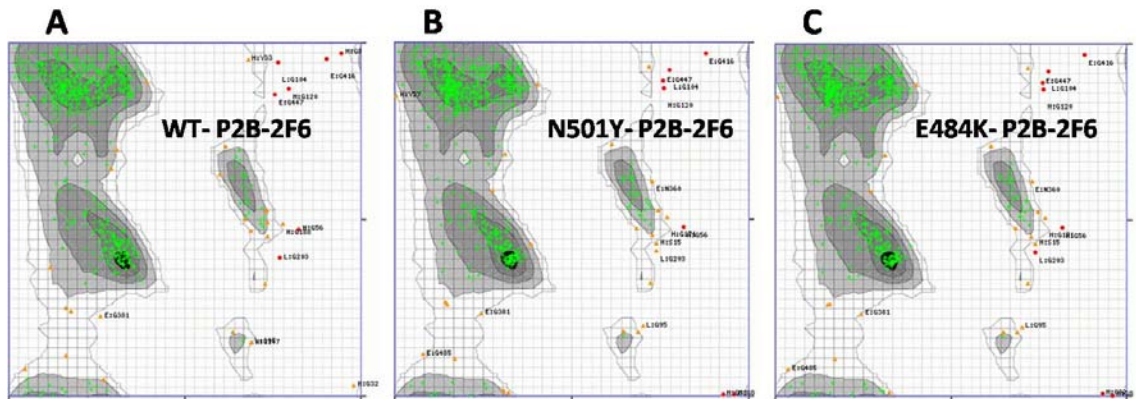


Figure S4

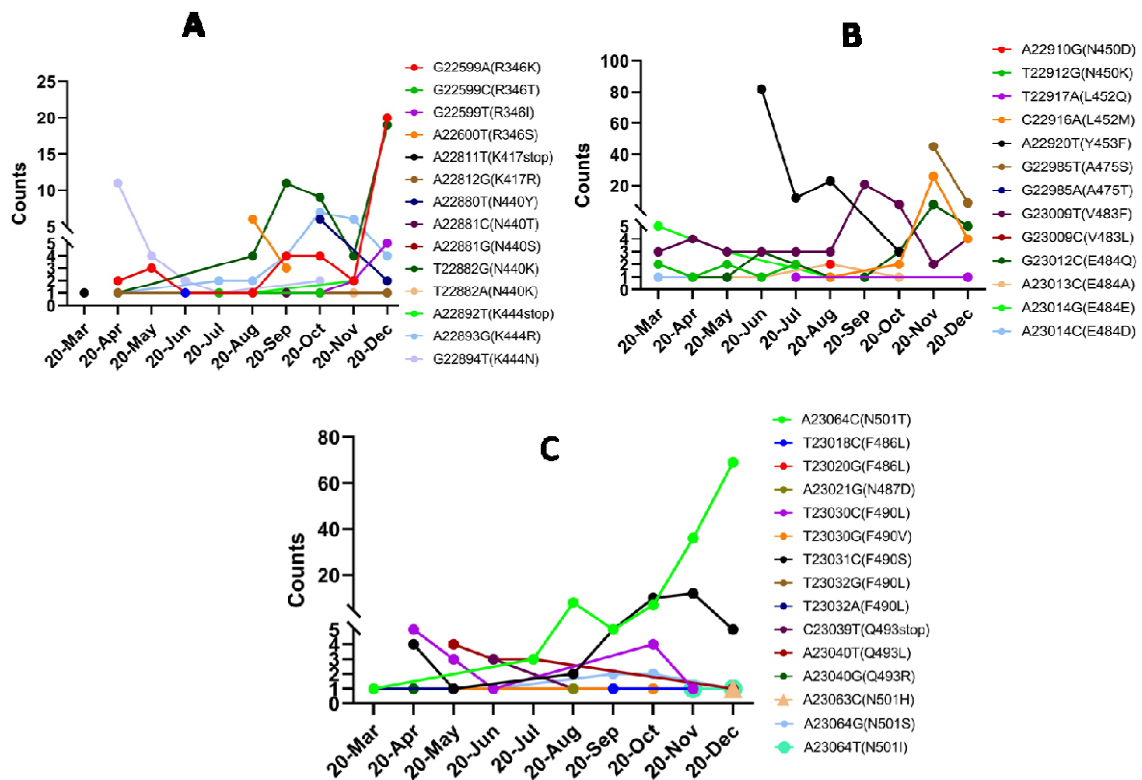




Figure S5

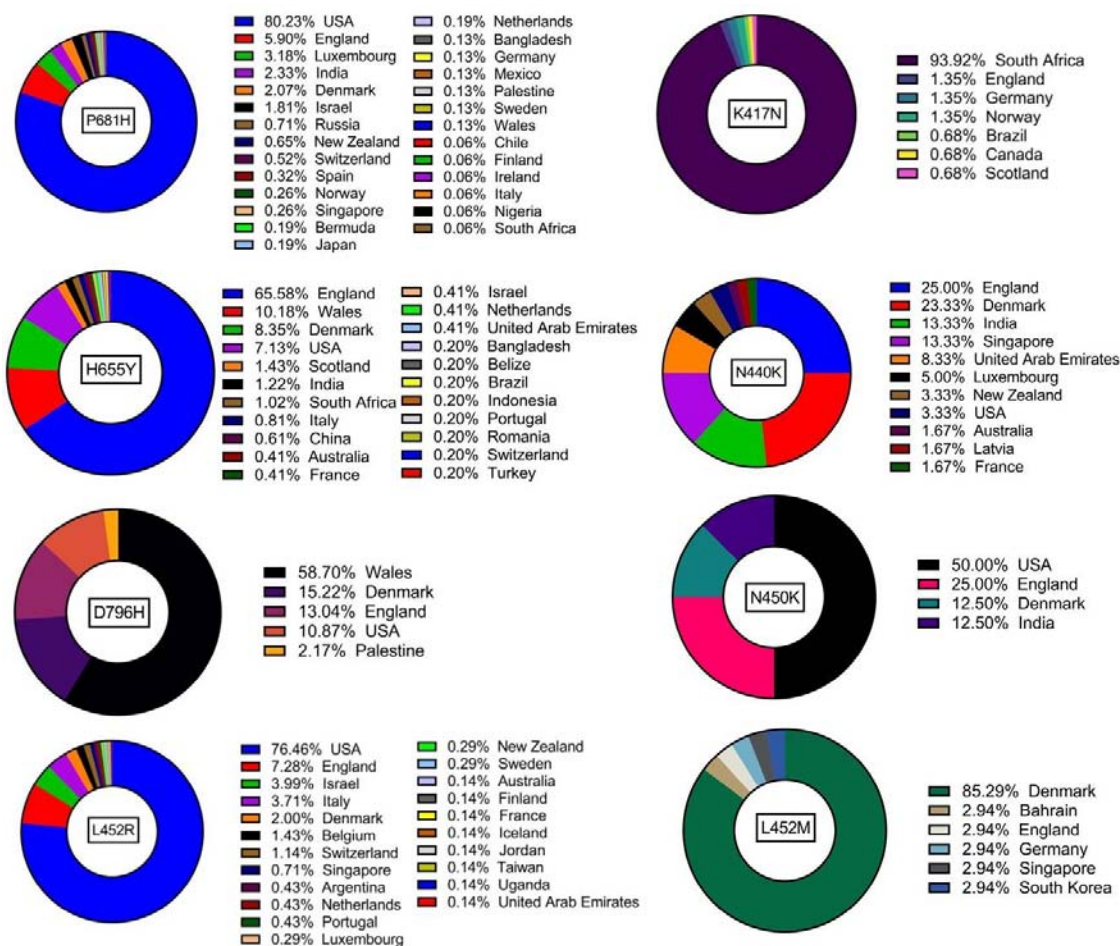
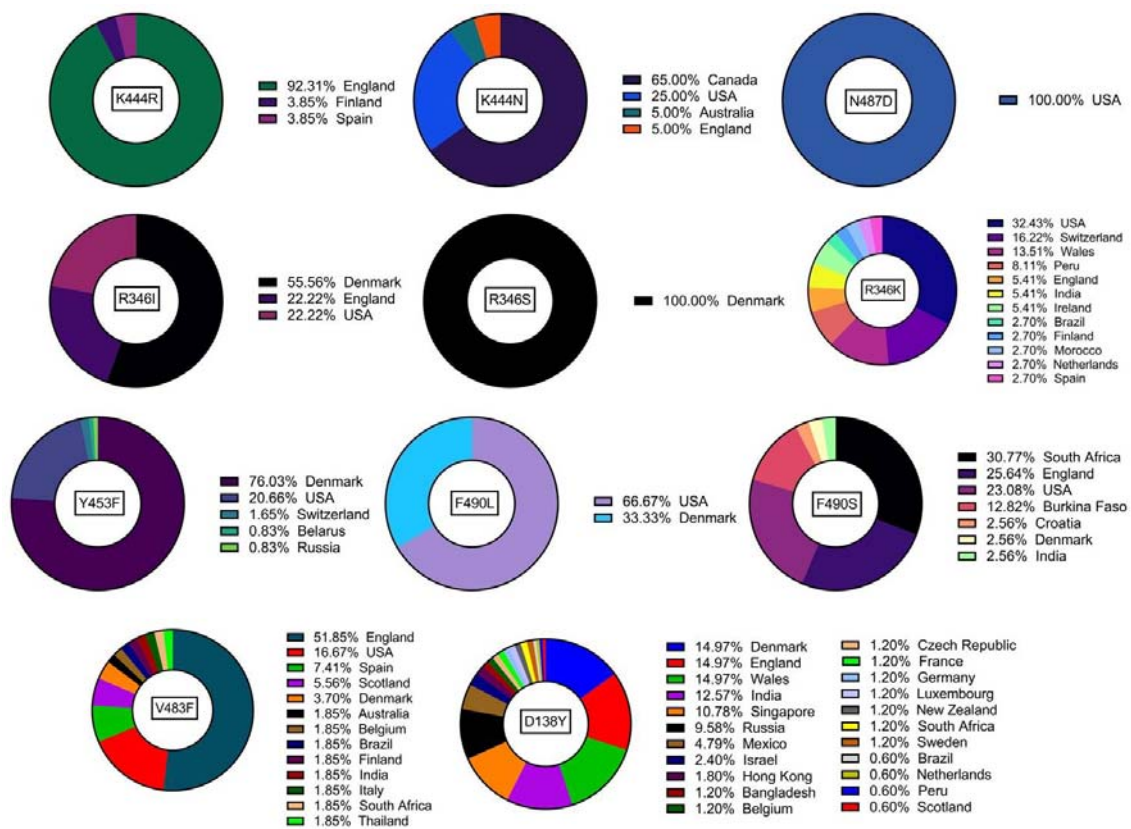


Figure S6



**Table 1:** Comparison of non-synonymous mutations identified in emerging SARS-COV-2 strains.

Gene	B1.1.7 (UK)	B1.351 (South Africa)	B1.1.28 (Brazil)	B1.1.28 (outside Brazil)	P1 Manaus (Brazil)
ORF1a	T1001I A1708D I2230T		L3468V L3930F	S1188L K1795Q	S1188L K1795Q
ORF1b				E1264D	E1264D
Spike	69-70HV del  144Y del         P681H  T716I S982A	L18F  D80A  D215G L242-244 del R246I K417N <b>E484K</b> <b>N501Y</b>  <i>D614G</i>  A701V	<b>E484K</b>	L18F T20N P26S  D138Y  R190S  K417T <b>E484K</b> <b>N501Y</b>  H655Y  T1027I	L18F T20N P26S  D138Y  R190S  K417T <b>E484K</b> <b>N501Y</b>  H655Y  T1027I

	D1118H		V1176F	V1176F	
ORF3a				S253P	
ORF8	Q27stop R52I Y73C			E92K	E92K
N	D3L S23F		A119S		P80R
ORF9b				Q77E	

**Table 2:** Inter and intra-molecular H-bond distances in the human ACE2 receptor in complex with spike protein RBD variants.

Interacting atoms <sup>a</sup>	N501 <sup>b</sup> Distance (Å)	N501Y <sup>c</sup> Distance (Å)
(hACE2) Y41OH-N501OD1(RBD)	3.43	-
(hACE2) K353-Y501OH (RBD)	-	2.79
(hACE2) Q42OE1-Q498NE2 (RBD)	3.38	3.06
(hACE2) Q42NE2-Y449OH (RBD)	2.79	2.8
(hACE2) D38OD2-Y449OH (RBD)	2.69	3.01
Q496O (O)-N501ND2 (RBD) <sup>a</sup>	3.4	-
Q496O (O)-Y501OH (RBD)	-	2.5
(hACE2)D38OD1-K353NZ (hACE2)	3.17	2.5

<sup>a</sup> Atoms in bracket represent peptide backbone amide oxygen or nitrogen.

<sup>b</sup> Based on PDB ID: 6M0J (Lan J, et al. 2020)

<sup>c</sup> Based on computationally generated structure through MutaBind2 server in this work.

**Table 3:** Inter-atomic distances at the interface of complex between P2B-2F6 neutralizing antibody and spike protein RBD variants

Interacting atoms <sup>a</sup>	Nature of interaction	E484 Distance (Å) <sup>b</sup>	E484K Distance (Å) <sup>c</sup>
(nAb-L) N33ND2-E484(O) (RBD) <sup>d</sup>	H-bond	2.5	3.02
(nAb-L) N33ND2-K484(O) (RBD)	H-bond	-	
(nAb-L) Y34OH-E484OE2 (RBD)	H-bond	2.47	-
(nAb-L) Y34OH-K484NZ (RBD)		-	7.07
(RBD) G485(N)- K484NZ (RBD)	H-bond	-	2.74
(nAb-H) R112NH1-E484OE2 (RBD)	Salt bridge	3.57	-
(nAb-H) R112NH1-E484OE1 (RBD)	Salt bridge	3.83	-
(nAb-H) Y27OH-G447(O) (RBD)	H-bond	2.93	3.21
(nAb-H) H54NE2-N450OD1 (RBD)	H-bond	2.83	3.01
(nAb-H) S30(O)-N450ND2 (RBD)	H-bond	3.28	3.28
(nAb-H) S31(O)-Y449(N) (RBD)	H-bond	3.11	2.88
(nAb-H) V106CG1-F490CE2 (RBD)	Van der waals	4.38	4.06
(nAb-H) V105CG2-F490CZ (RBD)	Van der waals	3.79	3.85
(nAb-H) V106CA-F490CZ (RBD)	Van der waals	4.04	4.14
(nAb-H) V105C-F490CZ (RBD)	Van der waals	4.05	4.02

<sup>a</sup> Light and heavy chains of the neutralizing antibody P2B-2F6 indicated by nAb-L and nAb-H, respectively.

<sup>b</sup> Based on PDB ID: 6M0J (Ju B, et al. 2020)

<sup>c</sup> Based on computationally generated structure through MutaBind2 server in this work.

<sup>d</sup> Atoms in bracket represent peptide backbone amide oxygen or nitrogen

**Table 4:** False negative rates for SARS-COV-2 RT-PCR test reported for population-based screening

Sample number <sup>a</sup>	Sample type	RT-PCR targets/kits	False negative (%)	Basis of false negative calling
100001 (Kanji JN et al, 2021)	Nasopharyngeal Oropharyngeal Deep nasal-turbinate swabs Endotracheal-aspirates Bronchoalveolar lavages	Laboratory designed testing Kit. Target genes: E and RdRp	9.3%	Two repeat RT-PCR test with target genes (E, N1 and N2). Mean time between two RT-PCR tests is 6.1 days.
20912 (Long DR, et al 2020)	Nasopharyngeal	RT-PCR was performed targeting the N, RdRp, S, and E and ORF1ab	4%	Repeat RT-PCR test withing 7 days. Mean interval between last negative and then a positive test is 4.0 days (SD-2.0 days).
1014 (Ai T, et al, 2020)	Throat swabs	TaqMan one step RT-PCR kits from shanghai huirui biotechnology Co Ltd, Target genes are orf1ab, N and E genes.	2%	Reconfirmed positive on basis of Chest CT report. Median time between chest CT and RT-PCR was 1 days (0-7 days range)
2021 (Baron R C, et al, 2020)	Nasopharyngeal and oropharyngeal swabs	Not mentioned	13%	Seropositivity

<sup>a</sup> References mentioned in bracket.



**Table 5:** False negative rates for SARS-COV-2 RT-PCR test reported for symptom-based testing

Sample number <sup>a</sup>	Sample type	RT-PCR targets/kits	False negative (%)	Basis of false negative calling
82 (He JL, et al. 2020)	Nasopharyngeal, oropharyngeal, Endotracheal aspirates Bronchoalveolar lavage	BGI genomes (Shenzhen, china) Target gene: - RdRp and S-protein	21%	Retested positive by RT-PCR after the observational period. Median RT-PCR test interval = 2(1-8 range)
25 (Duan X, et al. 2020)	Nasopharyngeal and oropharyngeal	CDC approved kits	12%	Retested positive on the basis of RT-PCR and blood test.
262 (An J, et al. 2020)	Nasopharyngeal swabs, Anal swabs	GeneoDXCo.Ltd, Shanghai, China.	14.5%	Retested RT-PCR positive within 5-7 days after symptom onset.
204 (Long C, et al. 2020)	Pharyngeal, oral and nasal samples	Not mentioned	17%	Based on positive CT report.
167 (Xie X et al. 2020)	Mouth swab	Not mentioned	3%	False negative reports confirmed, on basis of positive CT.
81 (Fang Y, et al. 2020)	Throat swabs and sputum.	RT-PCR (Shanghai ZJ Bio-Tech, Shanghai, China. The target genes are orf1ab/N/E.	29%	Confirmed through CT-report. Average interval = 3±3 days.
121	Bronchoalveolar	Sansure	12%	Retested RT-PCR positive.

(Bernheim A, et al. 2020)	lavage Endotracheal aspirate Nasopharyngeal swab, Oropharyngeal swab	Biotech(Changshu, China) Shanghai Zhejiang Biotechnology (Shanghai, China), Da An Gene (Guangzhou, China) Target gene: - N/orf1ab		
21 (Chen G, et al. 2020)	Nasopharyngeal swab, Oropharyngeal swab	Not mentioned	33%	Retested RT-PCR positive.
696 (Besutti, G, et al. 2020)	Nasopharyngeal and oropharyngeal	Gene Finder covid19 plus real amp kit Target gene: - RdRp/E/ N gene	2%	Retested RT-PCR positive. Repeat RT-PCR in 15 days was conducted.
33 (Chen ZH, et al. 2020)	Not mentioned	Not mentioned	36%	Retested RT-PCR positive. Median interval = 2 days (average=2.7 days, Range=1-6 days)

<sup>a</sup> References mentioned in bracket.

**Table 6:** False negative rates for SARS-COV-2 RT-PCR test reported for hospital-based

Sample number <sup>a</sup>	Sample type	RT-PCR targets/ kits	False negative (%)	Basis of false negative calling
172 (Yuan J, et al. 2020)	Cloacal and nasopharyngeal swabs	Covid-19 RNA detection kit (Real time fluorescent PCR-method) Target genes are orf1ab and N gene.	14.5%	Retested positive through RT-PCR kit and CT reports.
85 (Li C, et al. 2020)	Not mentioned	Not mentioned	17.6%	Retested RT-PCR positive and on the basis of CT-report. Average interval between second test 17( $\pm$ 6.72)days
161 (Xie C, et al. 2020)	Nasal swabs	Not mentioned	36.4%	Retested Positive RT-PCR. Median=7.5 days (1-14)
610 (Li Y, et al. 2020)	Pharyngeal swabs	Shanghai Huirui Biotechnology Co. Ltd. The target genes are orf1ab, N and E genes.	30%	Retested RT-PCR positive.
55 (Fechner C, et al. 2020)	Oropharyngeal, nasopharyngeal, tracheal secretions and BALF	Not mentioned	2%	Retested positive through thoracic CT-report and RT-PCR positive when specimens collected from BALF

592 (Lan FY, et al. 2020)	Nasopharyngeal swab	MADPH, CDC approved kit, Roche, Cobas, SARS-CoV-2 kit Abbott real time sars-cov-2 Target gene orf1a/b and E gene	10%	Retested positive.	RT-PCR
34 (Chen HJ, et al. 2020)	Not mentioned	Not mentioned	29%	Retested positive.	RT-PCR
185 (Çinkoo ğlu A., et al, 2020)	Not mentioned	Not mentioned	5%	Retested positive. Mean interval between two tests = $1.7 \pm 0.8$ days	RT-PCR
234 (Dai H., et al, 2020)	Not mentioned	Not mentioned	3%	Retested positive	RT-PCR
4 (Lan L., et al, 2020)	Throat swabs	Bio Germ RT-PCR test kit. Target genes are orf1ab and N genes.	100%	All the 4 patients had positive RT-PCR.	

<sup>a</sup> References mentioned in bracket.