



## Article

# SARS-CoV-2 Spike Mutations, L452R, T478K, E484Q and P681R, in the Second Wave of COVID-19 in Maharashtra, India

Sarah Cherian <sup>1,†</sup>, Varsha Potdar <sup>1,†</sup>, Santosh Jadhav <sup>1</sup>, Pragya Yadav <sup>1</sup> , Nivedita Gupta <sup>2</sup>, Mousumi Das <sup>1</sup>, Partha Rakshit <sup>3</sup>, Sujeet Singh <sup>3</sup>, Priya Abraham <sup>1,\*</sup>, Samiran Panda <sup>2</sup> and NIC Team <sup>‡</sup>

<sup>1</sup> ICMR-National Institute of Virology, Pune 411001, India; sarahcherian100@gmail.com (S.C.); potdarvarsha9@gmail.com (V.P.); smjnv@gmail.com (S.J.); hellopragya22@gmail.com (P.Y.); mosu39@gmail.com (M.D.)

<sup>2</sup> Indian Council of Medical Research, New Delhi 110029, India; drguptanivedita@gmail.com (N.G.); samiranpanda.hq@icmr.gov.in (S.P.)

<sup>3</sup> National Centre for Disease Control, New Delhi 110054, India; partho\_rakshit@yahoo.com (P.R.); sujeet647@gmail.com (S.S.)

\* Correspondence: priya.abraham@icmr.gov.in

† Equal authorship.

‡ Prachi Jagtap, Bhagyashri Kasabe, Ujjaini Shah, Tripti Sanjeev, Gayatri Divekar, Kalpita Korabu, Sunil Shelkande, Pooja Shinde, Sayed Zakiuddin, Veena Vipat, Sheetal Jadhav, Krutika Iyengar, Vinita Malik, Sonali Bhorekar, Abhinendra Kumar, Rima Sahay, Anita Shete, Manoharlal Choudhary from National Influenza Center (NIC).



**Citation:** Cherian, S.; Potdar, V.; Jadhav, S.; Yadav, P.; Gupta, N.; Das, M.; Rakshit, P.; Singh, S.; Abraham, P.; Panda, S.; et al. SARS-CoV-2 Spike Mutations, L452R, T478K, E484Q and P681R, in the Second Wave of COVID-19 in Maharashtra, India. *Microorganisms* **2021**, *9*, 1542. <https://doi.org/10.3390/microorganisms9071542>

Academic Editor: Mark E. Obrenovich

Received: 1 May 2021

Accepted: 1 July 2021

Published: 20 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** As the global severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic expands, genomic epidemiology and whole genome sequencing are being used to investigate its transmission and evolution. Against the backdrop of the global emergence of “variants of concern” (VOCs) during December 2020 and an upsurge in a state in the western part of India since January 2021, whole genome sequencing and analysis of spike protein mutations using sequence and structural approaches were undertaken to identify possible new variants and gauge the fitness of the current circulating strains. Phylogenetic analysis revealed that newly identified lineages B.1.617.1 and B.1.617.2 were predominantly circulating. The signature mutations possessed by these strains were L452R, T478K, E484Q, D614G and P681R in the spike protein, including within the receptor-binding domain (RBD). Of these, the mutations at residue positions 452, 484 and 681 have been reported in other globally circulating lineages. The structural analysis of RBD mutations L452R, T478K and E484Q revealed that these may possibly result in increased ACE2 binding while P681R in the furin cleavage site could increase the rate of S1-S2 cleavage, resulting in better transmissibility. The two RBD mutations, L452R and E484Q, indicated decreased binding to select monoclonal antibodies (mAbs) and may affect their neutralization potential. Further in vitro/in vivo studies would help confirm the phenotypic changes of the mutant strains. Overall, the study revealed that the newly emerged variants were responsible for the second wave of COVID-19 in Maharashtra. Lineage B.1.617.2 has been designated as a VOC delta and B.1.617.1 as a variant of interest kappa, and they are being widely reported in the rest of the country as well as globally. Continuous monitoring of these and emerging variants in India is essential.

**Keywords:** SARS-CoV-2; India; Maharashtra; evolution; second wave; whole genomes; B.1.617.1; B.1.617.2; modeling

## 1. Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) surface spike (S) protein mediates entry into host cells by binding to the host receptor angiotensin-converting enzyme 2 (ACE2) via its receptor-binding domain (RBD). Crystal structures of SARS-CoV-2 S protein or its RBD complexed with ACE2 from different hosts reveal that the RBD contains a core and a receptor-binding motif (RBM) which forms contacts with

ACE2 [1]. A number of naturally selected mutations in the RBM have been shown to affect infectivity, human-to-human transmission, pathogenesis and immune escape [2].

As the global SARS-CoV-2 pandemic expands, genomic epidemiology and whole genome sequencing are being used to investigate the transmission and evolution. The expansion of the genome-wide diversity resulted in delineation of the viral strains into clades, lineages and sub-lineages. Clades G/GH/GR/GV/GRY as per the Global Initiative on Sharing All Influenza Data (GISAID) database (<https://www.gisaid.org/>, Accessed on 14 June 2021) [3] possess a common mutation D614G in the spike protein. The mutation was shown to result in significantly higher human host infectivity and better transmission efficiency for the virus [4,5]. Three recently emerged “variants of concern” (VOCs) of SARS-CoV-2 are GRY (formerly GR/501Y.V1)/B.1.1.7 (alpha), GH/501Y.V2/B.1.351 (beta) and GR/501Y.V3/P.1 (gamma) [6]. These variants are known to possess multiple mutations across the genome, including several in the S protein and its RBD, such as N501Y, E484K and K417N/T [5–8]. Multiple SARS-CoV-2 variants are now seen to be circulating globally.

The potential introduction and consequence of these emerging variants in the country is vital to support the public health response. The National Influenza Centre at the ICMR-National Institute of Virology, Pune, as an apex laboratory of the Indian Council of Medical Research (ICMR), has been continuously involved in SARS-CoV-2 diagnostics and monitoring the genomic evolution of this virus. In addition, the ICMR-NIV, Pune is also one among the ten national laboratories of the Indian SARS-CoV-2 Consortium on Genomics (INSACOG), Ministry of Health and Family Welfare, Government of India, catering to western India. Against the backdrop of the global emergence of VOCs and an upsurge in Maharashtra, a state in the western part of India, enhanced sequencing was undertaken. This was to identify possible new variants and analyze spike protein mutants, in particular to gauge the fitness of current circulating strains, using bioinformatics sequence and molecular structural approaches.

## 2. Materials and Methods

The state of Maharashtra (latitude 19°39' N, longitude 75°18' E) is located in west central India, in the northwestern part of the Indian subcontinent. Since the end of January 2021, a concentrated spurt in COVID-19 cases was noted in several districts of Maharashtra (Figure S1). As per the state government directives, samples from international travelers and 5% of surveillance samples from the positive cases with Ct value <25, including those from clusters, long haulers and mild/moderate/severe and deceased cases, were referred to the ICMR-NIV, Pune for whole genome sequencing. Nasopharyngeal swabs (n = 2987, 25 November 2020–30 May 2021) were processed for whole genome sequencing.

Briefly, nucleic acid extraction was performed using 280 µL of each sample in duplicate by a Qiagen viral RNA extraction protocol and quantified RNA was further processed for template preparation using the Ion Chef System. Purified template beads were submitted to meta transcriptome next-generation sequencing (NGS) in the Ion S5 platform (ThermoFisher Scientific, Waltham, MA, USA) using an Ion 540™ chip and the Ion Total RNA-Seq kit v2.0, as per the manufacturer’s protocol (ThermoFisher Scientific, Waltham, MA, USA) [9]. Sequence data were processed using Torrent Suite Software (TSS) v5.10.1 (ThermoFisher Scientific, Waltham, MA, USA). Coverage analysis plugins were utilized to generate coverage analysis report for each of the samples. Reference-based reads gathering and assembly were performed for all the samples using Iterative Refinement Meta-Assembler (IRMA) [10]. A subset of the samples was also sequenced on an Illumina machine and analyzed using CLC Genomics Workbench version 20 (CLC, Qiagen, Hilden, Germany) as described elsewhere [11]. Of 2204 whole genomes obtained, 1791 were considered for different analyses based on coverage of ≥93% of the genome. Both eastern (n = 829) and western (n = 962) districts of the state were well represented. For each whole genome sequence, the GISAID clade assignment was done using CoVsurver: Mutation Analysis of hCoV-19 (<https://www.gisaid.org/epiflu-applications/covsurver-mutations-app/>, Accessed on 1 June 2021). For lineage assignment, a web application, i.e., Phylogenetic

Assignment of Named Global Outbreak LINEages (PangoLIN) COVID-19 Lineage Assigner (<https://pangolin.cog-uk.io/>, Accessed on 1 June 2021), was implemented [8]. The 1791 genomes were aligned using MAFFT v.7.450 [12] and a neighbor-joining phylogenetic tree was constructed using MEGA V.6 [13], employing the composite likelihood as the substitution model and 1000 bootstrap replications. The gene-wise top mutations in the predominant lineages were depicted using heat maps based on the calculations of the mutational frequencies.

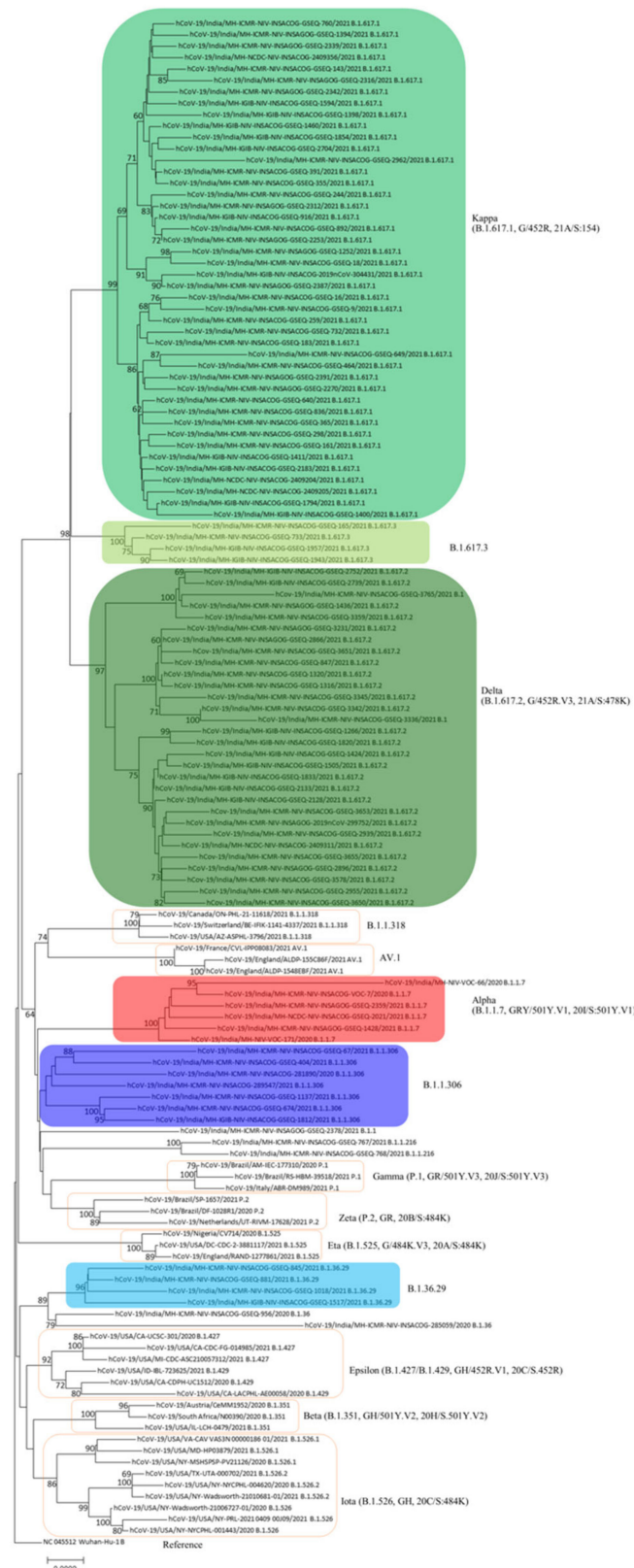
For further structural characterization of the S protein mutations, the crystal structure of SARS-CoV-2 S glycoprotein complexed with ACE2 was obtained from the protein data bank (PDB ID: 7A98 [14]). The top mutations in the S protein based on the heatmap were mapped using Biovia Discovery studio visualizer 2020. To assess the effect of noted RBD mutations on ACE2 binding, the crystal structure of the SARS-CoV-2 spike RBD domain complexed with ACE2 (PDB ID: 6LZG [15]) was used. For assessment of the noted mutations on binding to neutralizing antibodies, the SARS-CoV-2 spike RBD domain complexed with two selected mAbs REGN10933/P2B-2F6 were retrieved (PDB ID: 6XDG; resolution 3.90 Å and 7BWJ; resolution 2.65 Å, respectively) [16,17]. Point mutations were carried out using Biovia Discovery studio visualizer 2020 and the structures of the complexes were subjected to energy minimization using the macro model tool in Schrodinger 2020 using default parameters. The molecular interactions between the RBD–ACE2 interface, within the RBD and between the neutralizing mAbs-RBD were analyzed using a non-bonded interactions tool in Biovia Discovery studio visualizer 2020. The structure of the S protein of Wuhan-Hu-1 (Genbank ID: YP\_009724390.1) inclusive of the pre-cleavage motif PRRAR was modeled based on homology modeling using SWISS-MODEL (<https://swissmodel.expasy.org/>, Accessed on 31 May 2021) with the template as a SARS-CoV-2 spike protein trimer complex, 7CWU.pdb.

### 3. Results

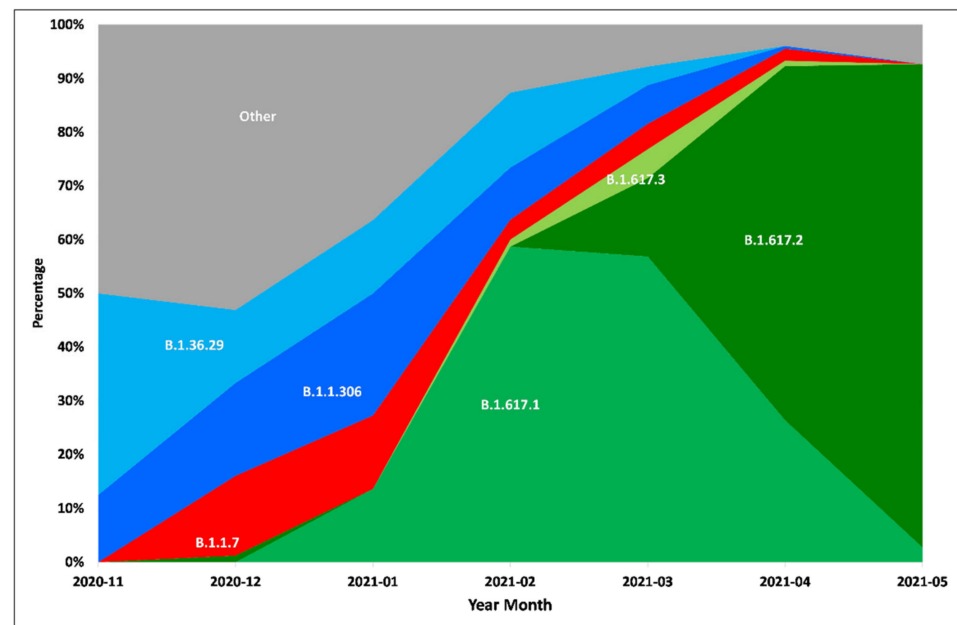
The genomic surveillance for the spurt in the COVID-19 cases in Maharashtra was carried out to identify the circulating lineages/VOCs and identify possible functionally significant mutations in the spike protein.

PangoLIN lineage classification of 1791 whole genomes revealed the presence of 40 lineages, with the majority within GISAID clades G/GH/GR/GRY/GV, evident from the D614G mutation in the S protein. Lineages B.1.617.1 ( $n = 779$ ), B.1.617.2 ( $n = 478$ ), B.1.1.306 ( $n = 116$ ), B.1.36.29 ( $n = 100$ ), B.1.1.7 ( $n = 75$ ), B.1.617.3 ( $n = 51$ ) were found to be the predominant lineages (Figure 1 and Figure S1). Chronologically, the earliest detected samples of lineage B.1.617.1, B.1.617.2 and B.1.617.3 were observed in January 2021, December 2020 and February 2021, respectively (Figure 2). The three new lineages, B.1.617.1, B.1.617.2 and B.1.617.3, could be linked to mutations specific to the spike region, along with ORF1a, ORF1b, ORF3a, M, ORF7a and N (Figure 3 and Figure S2).

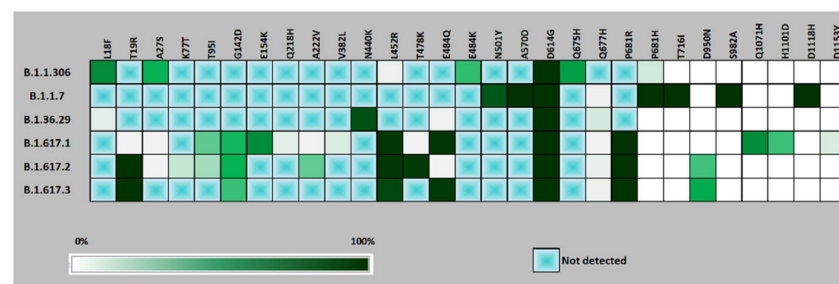
Among the new B.1.617 lineages, B.1.617.1 included the majority of the strains from eastern part of Maharashtra while B.1.617.2 also included sequences from major cities like Pune, Thane and Mumbai in the western part of the state. The mutations L452R and E484Q within the RBD were specific to lineage B.1.617.1 and B.1.617.3 while L452R and T478K were specific to lineage B.1.617.2. The lineage B.1.617.3 was characterized by mutations T19R and E484Q. Mutation G142D and P681R, within the spike but outside the RBD region, were common to all the three new lineages (Figure 3). P681R is noted in the S1-S2 furin cleavage site. Specific mutations such as T95I, H1101D and D1153Y were found in a proportion of B.1.617.1. Similarly, mutations K77T and A222V were found in a proportion of B.1.617.2. A synonymous mutation D111D was observed to be co-occurring with the RBD mutations L452R and E484Q in lineage B.1.617.1. Deletions 157/158 and deletions 119/120 with respect to lineage B.1.617.2 were noted.



**Figure 1.** Neighbor-joining tree of representative SARS-CoV-2 genomes depicting lineages in WHO label, PangoLIN, GISAID and Nextstrain. The sequences of this study are tagged as India/MH-ICMR-NIV and major lineages are in colored boxes. Additionally, included are representatives of the global WHO identified VOCs/VOIs/VUIs in unfilled boxes.



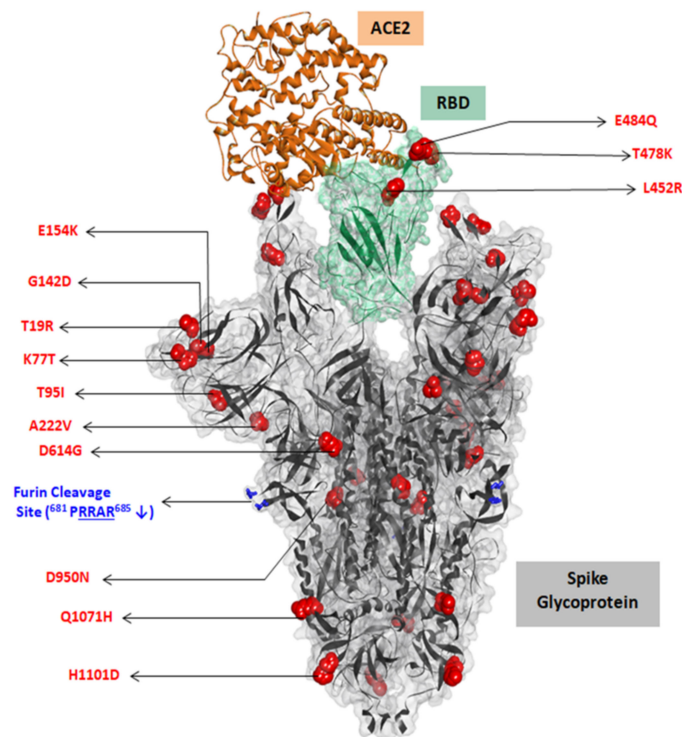
**Figure 2.** Temporal trend of major lineages in the districts of Maharashtra from November, 2020 to May, 2021.



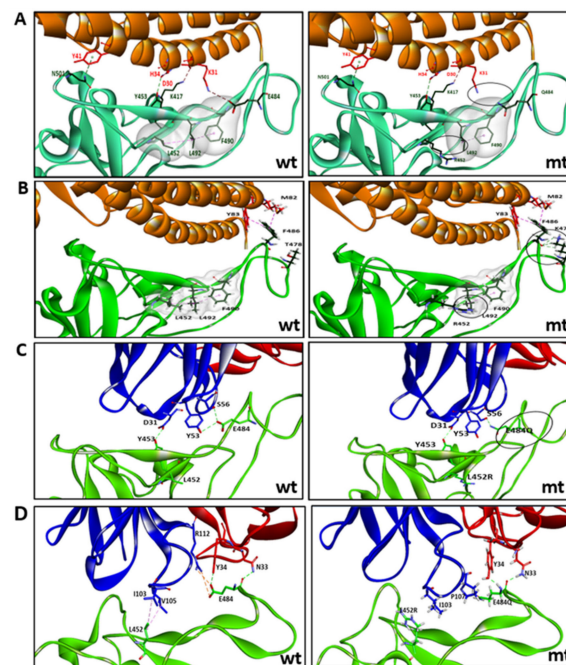
**Figure 3.** Heat map of major mutations in the spike protein of the predominant lineages from November, 2020 to May, 2021. A box with a green cross indicates that the specific mutation was absent.

The key mutations in the S protein are mapped on a furin-cleaved structure of the S protein (Figure 4). The structural implications of the RBD mutations, L452R and E484Q, as in lineages B.1.617.1 and B.1.617.3, were analyzed in terms of interaction with the ACE2 receptor and neutralizing antibodies that are known to have interactions with these residues (Figure S3). Similarly, the implications of RBD mutations, L452R and T478K, as in lineage B.1.617.2, were studied. The effect of the E484Q mutation is noted in terms of disruption in an electrostatic bond of the spike RBD residue E484 with K31 in the ACE2 interaction interface (Figure 5A, Supplementary Table S1). The intramolecular interactions in the wild strain indicate that the L452 residue is involved in a hydrophobic interaction with L492 which forms another hydrophobic contact with F490. These residue interactions form a hydrophobic patch on the surface of the RBD. The L452R mutation abolishes the hydrophobic interaction with L492 of the RBD. Estimation of the minimized energies of the wildtype and L452R occurring with the E484Q mutant structure of the RBD complexed with ACE2 showed energy values of  $-93732.305$  kcal/mol and  $-94543.180$  kcal/mol, respectively. In the case of the mutant RBD possessing L452R with T478K (Figure 5B), though no additional intermolecular contact was affected, three additional hydrogen bonds (H-bonds) were formed by K478 with F486. The intramolecular interaction L452–L492 associated with the L452R mutation was disrupted before. The minimum energy value of the RBD–ACE2 complex was  $94,751.367$  kcal/mol.





**Figure 4.** Mapping of key mutations on the furin-cleaved crystal structure of SARS-CoV-2 spike glycoprotein (gray surface view) in complex with ACE2 (solid brown ribbon). RBD region shown in green.



**Figure 5.** (A) Key interactions between ACE2–RBD involving mutations L452R and E484Q in the RBD, as in lineages B.1.617.1 and B.1.617.3. (B) Key interactions between ACE2–RBD involving mutations L452R and T478K in the RBD, as in lineage B.1.617.2. (C) Interactions between RBD–mAb REGN10933. (D) Interactions between RBD–mAb P2B-2F6. wt, corresponds to wildtype strain and mt, mutant strain. In (A) and (B) are the intra-molecular contacts in a hydrophobic patch of the RBD region (surface displayed in gray). In (C) and (D), blue represents heavy chain and red represents light chain.

The mutations L452R and E484Q are seen to disrupt the interfacial interactions of the spike RBD with specific neutralizing antibodies (Figure 5C,D; Supplementary Table S2). The heavy chain of monoclonal antibody REGN10933 interacts with the RBD by making two H-bonds between E484 of the RBD and Y53 and S56 of the antibody. Residue Y453 of the RBD makes an H-bond with D31 of the heavy chain of the antibody. The RBD mutation E484Q disrupts the two H-bonds with S56 and Y53 (Figure 5C). Residues L452 and E484 possess direct contacts with the mAb P2B-2F6 (Figure 5D). The intermolecular interactions of the wildtype complex indicate that L452 is involved in hydrophobic interactions with residues I103 and V105 of the heavy chain of the antibody. Residue E484 forms H-bonds with N33 and Y34 of the light chain as well as a salt bridge and H-bond interaction with the R112 side chain of the light chain. The L452R mutation breaks the hydrophobic interactions with both residues I103 and V105 and also disrupts the H-bond and electrostatic interaction with R112.

Modeling the S protein with the pre-cleavage motif PRRAR showed the exposure of R682 and R685 (Figure S4). Induction of a P681R point mutation in the modeled structure further revealed that the side chain of R681 could facilitate additional interactions with furin.

#### 4. Discussion

In addition to the three global VOCs, the global variants of interest (VOIs) such as GH/452R.V1/ B.1.427/B.1.429 (epsilon), GR/P.2 (zeta), G/484K.V3/B.1.525 (eta), GR/P.3 (theta) and GH/B.1.526 (iota) have been identified by the WHO [6]. The potential consequences of emerging variants are increased transmissibility, increased pathogenicity and the ability to escape natural or vaccine-induced immunity [18,19]. With this background, we sequenced and analyzed the whole genomes of SARS-CoV-2 from different districts of Maharashtra where a surge in COVID-19 activity was noted since the end of January 2021 after a gap of almost four months.

Phylogenetic and sequence analyses revealed that only a small proportion of sequences, mostly in the month of December 2020, were detected as the VOC “alpha”. The deletions 69/70 and 144/145 specific to this lineage were noted in these strains. On the other hand, the B.1.617.1 lineage possessing common signature mutations L452R, E484Q and P681R in the spike protein could be linked to the surge of cases in February 2021 in eastern Maharashtra. In addition, districts in western Maharashtra, such as Pune, Mumbai, Thane and Nashik, showed the presence of multiple lineages in circulation in comparison to the dominance of lineage B.1.617.1 in eastern Maharashtra (Figure S1). From March 2021, lineage B.1.617.1 declined and was overtaken by B.1.617.2. The epidemiological evidence with respect to the surge being noted in the eastern districts suggests that B.1.617.1 has its epicenter in that part of Maharashtra. Further, movements associated with social gatherings and local body elections most likely led to the spread to the western parts of the state. This is where major highly populous metropolitan cities are located.

A comparison of the lineages circulating in India during the study period from November 2020 to May 2021 (Figure S5A), revealed that from among the new lineages, only B.1.617.1 and B.1.617.2 were the predominant ones in circulation. Analyzing retrospectively, there were single sequences found to be submitted to GISAID of lineage B.1.617.1 in December 2020 from Maharashtra (district Nagpur) and Odisha. Similarly, B.1.617.2 was noted in September 2020 from the state of Madhya Pradesh, in November 2020 from Uttar Pradesh and in December in Bihar and Maharashtra. The proportion of B.1.617.1 and B.1.617.2 in Maharashtra from February–March 2021 was found to be about 55–60% and 10–60%, respectively, as compared to 5 to 15% and 3–10% in the rest of the country (Figure S5B). Less sequencing before the study period may have resulted in missing out on earlier detections of these lineages. Overall, the observations may suggest the dispersal of the new lineages from Maharashtra to the other states in India. The early detection of the variants in the states of Uttar Pradesh, Bihar and Odisha could be attributed to the movement of migrant workers from the state of Maharashtra. In addition, in India during

the period from November 2020 to May 2021, there were other lineages, such as B.1, B.1.1, B.1.36, B.1.351, B.1.618 and B.1.1.216, that were not common in the state of Maharashtra. Our earlier studies [9,20] that were based on genome sequences from India from January to August/September 2020 had demonstrated that the dominant lineages were B.1.113, B.1.1.32, B.1.1.8, B.1.80, B.4 and B.6. This reflects the diversity of lineages in the country and also the switches in the dominant lineages.

Notably, D111D was found to be associated with the signature mutations of lineage B.1.1.617.1. The co-occurrence of synonymous mutations with the non-synonymous mutations observed is interesting and generally being reported [21,22].

The structural analysis of the effect of RBD mutations L452R and E484Q towards ACE2 binding revealed a decrease in intramolecular and intermolecular contacts with respect to the wildtype. However, the hydrophobic L452 residue mutation to the hydrophilic 452R might help in interactions with water molecules and overall stabilization of the complex, as was reflected in the lower minimum energy of the mutant complex. The effect of the mutations L452R and T478K on ACE2 binding was also observed as enhanced stabilization of the RBD–ACE2 complex.

Another significant mutation, P681R, in the furin cleavage site resulted in enhancement of the basicity of the poly-basic stretch, and the likely facilitation of additional contacts with furin for S1–S2 cleavage. This could help in an increased rate of membrane fusion, internalization and thus better transmissibility. A recent study [23] revealed increased infectivity of the B.1.617 spike protein that could be attributed to L452R which itself caused a 3.5-fold increase in infectivity and, in combination with E484Q, caused a 3-fold increase. The P681R mutation caused a small increase in proteolytic processing that might have an effect on infectivity [24]. Observations on enhanced transmissibility of B.1.617.2 are also reported [23]. Further *in vitro* and/or *in vivo* studies to understand the phenotypic effects of the mutant strains would be vital for validating these observations.

Structural analysis further showed that the two RBD mutations L452R and E484Q may decrease the binding ability of REGN10933 and P2B-2F6 antibodies to the variant strains, compared to that in the wildtype strain. A recent report [25] revealed that the L452R mutation reduced or abolished neutralizing activity of 14 out of 35 RBD-specific mAbs including three clinical stage mAbs. Another recent study [26] has demonstrated that the mutation L452R can escape from human leukocyte antigen (HLA) 24-restricted cellular immunity and can also increase viral infectivity, potentially promoting viral replication. The combination of the two RBD mutations, L452R and E484Q, noted in this study could affect the neutralization of the select mAbs. The neutralizing potential of a wider range of mAbs against the strains of the B.1.617 lineages needs to be assessed. Recent studies have investigated the level of neutralization efficacy of sera from vaccinees of different vaccine formulations and those who have experienced natural infection. The neutralization of B.1.617.1 with sera of BBV152 (Covaxin) vaccinees showed almost two-fold reduction in neutralization compared to a B.1 “G” clade strain [27]. Even more reduction in neutralizing titer with sera of ChAdOx1 2 (Covishield)-vaccinated individuals was observed against the B.1.617.1 variant [28]. Against the B.1.617.2 variant, a 4.6-fold and 2.7-fold reduction in the neutralization titer was observed in comparison to the B.1 variant with sera of COVID-19 recovered cases and Covaxin vaccinees, respectively [29]. Bernal et al. [30], have also reported the effectiveness of COVID-19 vaccines against the B.1.617.2 variant. A summary of the major spike protein mutations defining the B.1.617 lineages and their proposed effects [31–37] is provided in Supplementary Table S3.

Mutations at both the residue positions 452 and 484 individually have been reported previously. L452R has been noted earlier in lineages B.1.427 and B.1.429 while the E484K mutation is common to the two VOCs beta and gamma and two VUIs, zeta and eta. E484Q has also been reported in several sequences in the GISAID. P681H is one of the mutations in the VIC alpha, while P681R is one of the mutations in the lineage A.23.1, which is identified as a “variant under monitoring” (<https://www.ecdc.europa.eu/en/COVID-19/variants-concern>, Accessed on 14 June 2021). The new lineages B.1.617.1 and B.1.617.3 in this study



are thus a unique combination of spike mutations L452R, E484Q and P681R. Between these two lineages, it was observed that B.1.617.1 spread more rampantly. The T478K mutation in the spike protein in lineage B.1.617.2 has been seen in Mexican variant B.1.1.222. This mutation or its combination with the other mutations in the genome appears to give a greater fitness advantage to the B.1.617.2 lineage as this lineage took over from B.1.671.1, and became dominant in Maharashtra and India. As of June 12, 2021, the worldwide prevalences of B.1.617.1 and B.1.617.3 are <0.5% while the prevalence of B.1.617.2 is 2% (outbreak.info). As per GISAID submissions, the B.1.617.2 lineage has been reported from 64 countries including the UK, USA, Germany, Singapore, Japan, Australia, etc. The global spread of the lineage resulted in its designation as a VOC delta. Lineage B.1.617.1 has been assigned as a VOI kappa [6].

To summarize, the study investigated the S protein mutations associated with the rise in COVID-19 cases in Maharashtra observed since the month of February 2021. The continuous increase in positivity could be attributed to signature mutations in the spike protein and functionally significant co-occurring mutations. The investigations related to patterns of circulation of the new variant lineages during the COVID-19 second wave in India need to be continued for their public health impact and association with vaccine breakthroughs, both in the case of partially and fully vaccinated individuals and reinfections.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/microorganisms9071542/s1>, Figure S1: The distribution of predominant PangoLIN lineages in the labelled districts of Maharashtra. The size of the pie charts is proportional to the number of genome sequences from the particular district. Figure S2: Heat map of major mutations in the different proteins of the predominant lineages from November, 2020 to May, 2021. Figure S3: Structural complexes of RBD-ACE2 and RBD-Fab. A. RBD—ACE2 (PDB Id-6LZG), B. RBD—mAb REGN10933(PDB Id-6XDG) C. RBD—mAb P2B-2F6(7bwj). Figure S4: Modeling of the SARS-CoV-2 spike protein (with the furin cleavage motif) and effect of P681R mutation (A) Complete spike protein with the residues in the furin cleavage site labelled (B) Wild-type strain with P681 (C) Mutant strain with P681R. Figure S5: Temporal trend of major lineages (top twelve) in India based on sequences from GISAID (As of June 6, 2021) from November, 2020 to May, 2021 (A) India inclusive of Maharashtra (n = 9495 ) (B) India excluding Maharashtra (n = 6949). Table S1: Intra and intermolecular interactions of the wild-type and mutant spike RBD of SARS-CoV-2 with ACE2. The bonds that are made bold in the wild-type indicate those that are disrupted in the mutant strain while the bonds that are bold in the mutant indicate those that are additional when compared to the wild-type strain. Mutant1 indicates the strain of lineage B.1.617.1 and mutant2 indicates a strain of lineage B.1.617.2. Table S2: Intermolecular interaction of neutralizing antibodies with wild-type and mutant spike RBD of SARS-CoV-2. Table S3: Major spike protein mutations defining the B.1.617 lineages and proposed effect on structural aspects.

**Author Contributions:** Conceptualization, S.C., V.P., P.A.; methodology, S.C., V.P., P.Y., N.G.; software, S.J.; formal analysis, S.C., V.P.; investigation, M.D., P.R.; writing—original draft preparation, S.C., V.P.; writing—review and editing, S.C., V.P., P.A.; supervision, S.C., V.P.; project administration, P.A., N.G., S.S., S.P.; funding acquisition, P.A., N.G., S.P. NIC Team author contributions: Softwares, P.J., B.K.; investigation, U.S., T.S., G.D., K.K., S.S., P.S., S.Z., V.V., S.J., K.I., V.M., S.B., A.K., R.S., A.S., M.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** The study was supported by the Indian Council of Medical Research, Department of Health Research, Ministry of Health and Family welfare, Government of India.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study, considering that the study was conducted on referred samples as part of the COVID-19 pandemic investigations.

**Informed Consent Statement:** Patient consent was waived considering that the study was conducted on referred samples as part of the COVID-19 pandemic investigations.

**Data Availability Statement:** Genomes sequenced as a part of this study have been deposited in GISAID, with INSACOG tags.

**Acknowledgments:** The authors acknowledge Indian SARS-CoV-2 Consortium on Genomics (IN-SACOG), Ministry of Health and Family Welfare, Government of India for their support. The authors are grateful to State Integrated Disease Surveillance Program (IDSP) for coordination of samples. We also acknowledge the cohesive efforts of NIC staff for sample processing and diagnostics. Thanks are due to Bioinformatics Group members for inputs during analyses.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221–224. [CrossRef]
- Wan, Y.; Shang, J.; Graham, R.; Baric, R.S.; Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: An Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* **2020**, *94*, e00127–20. [CrossRef] [PubMed]
- Elbe, S.; Buckland-Merrett, G. Data, disease and diplomacy: GISAID’s innovative contribution to global health. *Glob. Chall.* **2017**, *1*, 33–46. [CrossRef] [PubMed]
- Korber, B.; Fischer, W.M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E.E.; Bhattacharya, T.; Foley, B.; et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **2020**, *182*, 812–827.e19. [CrossRef]
- Volz, E.; Hill, V.; McCrone, J.T.; Price, A.; Jorgensen, D.; O’Toole, Á.; Southgate, J.; Johnson, R.; Jackson, B.; Nascimento, F.F.; et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. *Cell* **2021**, *184*, 64–75.e11. [CrossRef] [PubMed]
- World Health Organization. Tracking SARS-CoV-2 Variants. Available online: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/> (accessed on 12 June 2021).
- Leung, K.; Shum, M.H.; Leung, G.M.; Lam, T.T.; Wu, J.T. Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020. *Eurosurveillance* **2021**, *26*, 2002106. [CrossRef]
- Rambaut, A.; Holmes, E.C.; O’Toole, Á.; Hill, V.; McCrone, J.T.; Ruis, C.; du Plessis, L.; Pybus, O.G. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **2020**, *5*, 1403–1407. [CrossRef]
- Potdar, V.; Vipat, V.; Ramdasi, A.; Jadhav, S.; Pawar-Patil, J.; Walimbe, A.; Patil, S.S.; Choudhury, M.L.; Shastri, J.; Agrawal, S.; et al. Phylogenetic classification of the whole-genome sequences of SARS-CoV-2 from India & evolutionary trends. *Indian J. Med. Res.* **2021**, *153*, 166–174.
- Shepard, S.S.; Meno, S.; Bahl, J.; Wilson, M.M.; Barnes, J.; Neuhaus, E. Viral deep sequencing needs an adaptive approach: IRMA, the iterative refinement meta-assembler. *BMC Genom.* **2016**, *17*, 708.
- Yadav, P.D.; Potdar, V.A.; Choudhary, M.L.; Nyayanit, D.A.; Agrawal, M.; Jadhav, S.M.; Majumdar, T.D.; Shete-Aich, A.; Basu, A.; Abraham, P.; et al. Full-genome sequences of the first two SARS-CoV-2 viruses from India. *Indian J. Med. Res.* **2020**, *151*, 200–209.
- Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [CrossRef] [PubMed]
- Tamura, K.; Stecher, G.; Peterson, D.; Filipiński, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2727. [CrossRef]
- Benton, D.J.; Wrobel, A.G.; Xu, P.; Roustan, C.; Martin, S.R.; Rosenthal, P.B.; Skehel, J.J.; Gamblin, S.J. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* **2020**, *588*, 327–330. [CrossRef] [PubMed]
- Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, A.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.-Y.; et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, *181*, 894–904.e9. [CrossRef]
- Hansen, J.; Baum, A.; Pascal, K.E.; Russo, V.; Giordano, S.; Wloga, E.; Fulton, B.O.; Yan, Y.; Koon, K.; Patel, K.; et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science. Am. Assoc. Adv. Sci.* **2020**, *369*, 1010–1014.
- Ju, B.; Zhang, Q.; Ge, J.; Wang, R.; Sun, J.; Ge, X.; Yu, J.; Shan, S.; Zhou, B.; Song, S.; et al. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **2020**, *584*, 115–119. [CrossRef] [PubMed]
- Greaney, A.J.; Loes, A.N.; Crawford, K.H.D.; Starr, T.N.; Malone, K.D.; Chu, H.Y.; Bloom, J.D. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* **2021**, *29*, 463–476.e6. [CrossRef] [PubMed]
- Garcia-Beltran, W.F.; Lam, E.C.; St Denis, K.; Nitido, A.D.; Garcia, Z.H.; Hauser, B.M.; Feldman, J.; Pavlovic, M.N.; Gregory, D.J.; Poznansky, M.C.; et al. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. *Cell* **2021**, *184*, 2372–2383.e9. [CrossRef]
- Yadav, P.D.; Nyayanit, D.A.; Majumdar, T.; Patil, S.; Kaur, H.; Gupta, N.; Shete, A.M.; Pandit, P.; Kumar, A.; Aggarwal, N.; et al. An Epidemiological Analysis of SARS-CoV-2 Genomic Sequences from Different Regions of India. *Viruses* **2021**, *13*, 925. [CrossRef]
- Sarkar, R.; Mitra, S.; Chandra, P.; Saha, P.; Banerjee, A.; Dutta, S.; Chawla-Sarkar, M. Comprehensive analysis of genomic diversity of SARS-CoV-2 in different geographic regions of India: An endeavour to classify Indian SARS-CoV-2 strains on the basis of co-existing mutations. *Arch. Virol.* **2021**, *166*, 801–812. [CrossRef]
- Zeng, H.-L.; Dichio, V.; Horta, E.R.; Thorell, K.; Aurell, E. Global analysis of more than 50,000 SARS-CoV-2 genomes reveals epistasis between eight viral genes. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 31519–31526. [CrossRef] [PubMed]

23. England, P.H. SARS-CoV-2 Variants of Concern and Variants under Investigation in England Technical Briefing 10. 2021. Available online: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/984274/Variants\\_of\\_Concern\\_VOC\\_Technical\\_Briefing\\_10\\_England.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/984274/Variants_of_Concern_VOC_Technical_Briefing_10_England.pdf) (accessed on 14 June 2021).
24. Tada, T.; Zhou, H.; Dcosta, B.M.; Samanovic, M.I.; Mulligan, M.J.; Landau, N.R. The Spike Proteins of SARS-CoV-2 B.1.617 and B.1.618 Variants Identified in India Provide Partial Resistance to Vaccine-elicited and Therapeutic Monoclonal Antibodies. *bioRxiv* **2021**. [[CrossRef](#)]
25. McCallum, M.; Bassi, J.; De Marco, A.; Chen, A.; Walls, A.C.; Di Iulio, J.; Tortorici, M.A.; Navarro, M.-J.; Silacci-Fregni, C.; Saliba, C.; et al. SARS-CoV-2 immune evasion by variant B.1.427/B.1.429. *bioRxiv* **2021**. [[CrossRef](#)]
26. Motozono, C.; Toyoda, M.; Zahradnik, J.; Ikeda, T.; Saito, A.; Tan, T.S.; Ngare, I.; Nasser, H.; Kimura, I.; Uriu, K.; et al. An emerging SARS-CoV-2 mutant evading cellular immunity and increasing viral infectivity. *bioRxiv* **2021**. [[CrossRef](#)]
27. Yadav, P.D.; Sapkal, G.N.; Abraham, P.; Ella, R.; Deshpande, G.; Patil, D.Y.; Nyayanit, D.A.; Gupta, N.; Sahay, R.R.; Shete, A.M.; et al. Neutralization of variant under investigation B.1.617 with sera of BBV152 vaccinees. *Clin. Infect. Dis.* **2021**, ciab411. [[CrossRef](#)]
28. Yadav, P.D.; Sapkal, G.N.; Abraham, P.; Deshpande, G.; Nyayanit, D.A.; Patil, D.Y.; Gupta, N.; Sahay, R.R.; Shete, A.M.; Kumar, S.; et al. Neutralization potential of Covishield vaccinated individuals sera against B.1.617.1. *Clin. Infect. Dis.* **2021**, ciab483. [[CrossRef](#)]
29. Yadav, P.D.; Sapkal, G.N.; Ella, R.; Sahay, R.R.; Nyayanit, D.A.; Patil, D.Y.; Deshpande, G.; Shete, A.M.; Gupta, N.; Mohan, V.K.; et al. Neutralization against B.1.351 and B.1.617.2 with sera of COVID-19 recovered cases and vaccinees of BBV152. *bioRxiv* **2021**. [[CrossRef](#)]
30. Bernal, J.; Andrews, N.; Gower, C.; Gallagher, E.; Simmons, S.S.; Thelwall, S. Effectiveness of COVID-19 vaccines against the B.1.617.2 variant. *medRxiv* **2021**. [[CrossRef](#)]
31. Li, Q.; Wu, J.; Nie, J.; Zhang, L.; Hao, H.; Liu, S.; Zhao, C.; Zhang, Q.; Liu, H.; Nie, L.; et al. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell* **2020**, *182*, 1284–1294.e9. [[CrossRef](#)] [[PubMed](#)]
32. Chen, J.; Gao, K.; Wang, R.; Wei, G.-W. Revealing the threat of emerging SARS-CoV-2 mutations to antibody therapies. *bioRxiv* **2021**. [[CrossRef](#)]
33. Liu, Z.; VanBlargan, L.A.; Bloyet, L.-M.; Rothlauf, P.W.; Chen, R.E.; Stumpf, S.; Zhao, H.; Errico, J.M.; Errico, E.S.; Liebeskind, M.J.; et al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe*. **2021**, *29*, 477–488.e4. [[CrossRef](#)] [[PubMed](#)]
34. Deng, X.; Garcia-Knight, M.A.; Khalid, M.M.; Servellita, V.; Wang, C.; Morris, M.K.; Sotomayor-González, A.; Glasner, D.R.; Reyes, K.R.; Gliwa, A.S.; et al. Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. *Cell* **2021**, *184*, 3426–3437.e8. [[CrossRef](#)] [[PubMed](#)]
35. Tchesnokova, V.; Kulakesara, H.; Larson, L.; Bowers, V.; Rechkina, E.; Kisiela, D.; Sledneva, Y.; Choudhury, D.; Maslova, I.; Deng, K.; et al. Acquisition of the L452R mutation in the ACE2-binding interface of Spike protein triggers recent massive expansion of SARS-CoV-2 variants. *bioRxiv* **2021**. [[CrossRef](#)]
36. Hoffmann, M.; Hofmann-Winkler, H.; Krüger, N.; Kempf, A.; Nehlmeier, I.; Graichen, L.; Sidarovich, A.; Moldenhauer, A.-S.; Winkler, M.S.; Schulz, S.; et al. SARS-CoV-2 variant B.1.617 is resistant to Bamlanivimab and evades antibodies induced by infection and vaccination. *bioRxiv* **2021**. [[CrossRef](#)]
37. Di Giacomo, S.; Mercatelli, D.; Rakhimov, A.; Giorgi, F.M. Preliminary report on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Spike mutation T478K. *J. Med. Virol.* **2021**. [[CrossRef](#)] [[PubMed](#)]