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

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

Virology 565 (2022) 29–37

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology

Dynamics of SARS-CoV-2 variants of concern (VOC) in Bangladesh during the first half of 2021

Sultana Zahura Afrin [a], Md Taohidul Islam [b], Shyamal Kumar Paul [c], Nobumichi Kobayashi [d], Rokshana Parvin [e,*]

[a] Department of Microbiology, Sheikh Hasina Medical College, Tangail, Bangladesh
[b] Population Medicine and AMR Laboratory, Department of Medicine, Faculty of Veterinary Science, Bangladesh Agricultural University, Mymensingh, 2202, Bangladesh
[c] Department of Microbiology, Netrokona Medical College, Netrokona, Bangladesh
[d] Department of Hygiene, Sapporo Medical University School of Medicine, Sapporo, Japan
[e] Department of Pathology, Faculty of Veterinary Science, Bangladesh Agricultural University, Mymensingh, 2202, Bangladesh

ARTICLE INFO

Keywords:
SARS-CoV-2
Sequence
Phylogeny
Variant
Mutation
Bangladesh

ABSTRACT

Bangladesh is the second-worst-affected country in South Asia by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The aim of this study is to examine genome sequences from Bangladesh from January 2021 to June 2021 in order to monitor the SARS-CoV-2 VOC and the clades or lineages that are prevalent in the country. Within the study timeframe, at least eight Nextstrain clades were found: 20A, 20B, 20C, 20H (Beta, V2), 20I (Alpha, V1), 20 J (Gamma, V3), 21A (Delta), 21D (Eta), and six GISAID clades: four main (G, GH, GR, GRY) and two minors (GV, O) with an introduction of VOC B.1.1.7/Alpha, B.1.351/Beta and B.1.617.2/Delta. The introduction and recent occurrence of VOCs with substantial alterations in the receptor binding site of spike protein (K417 N, K417T, L452R, T478K, E484K, S494P, N501Y) are of particular importance. Specifically, VOC B.1.617.2/Delta has surpassed all prior VOCs in Bangladesh, posing a challenge to the existing disease management.

1. Introduction

For more than a year and half, the world has been dealing with the Covid-19 pandemic. Covid-19 is one of the most serious zoonotic diseases the world has ever seen, and the causative agent is severe acute respiratory syndrome virus type 2 (SARS-CoV-2) the important member of Coronavirus family. Coronaviruses are notable for breaching the 'species barrier' resulting in zoonotic and reverse zoonotic disease (Kumar et al., 2021). It has caused significant illness and mortality around the world since December 2019. SARS-CoV-2 is an enveloped, positive sense, single stranded RNA virus. Until now, the origin of the SARS-CoV-2 is unclear however, bat coronavirus RaTG13 is the closest relative of SARS-CoV-2, (Li, T. et al., 2020; Wu et al., 2020; Zhou et al., 2020). SARS-CoV-2 has accumulated numerous mutations since its origin due to its fast transmission among a large population in a short period of time (Li, X. et al., 2020). The majority of mutations have little or no influence on viral transmission or diseases progression (Chen et al., 2020). However, several mutations in the Spike (S) gene regions of SARS-CoV-2 genome are believed to alter the transmissibility pattern (Li, X. et al., 2020; Chen et al., 2020). As a result of these distinctive changes in S and few other genes, numerous variants have emerged. These variants have been a major concern as current vaccine strategies were developed on the prototype strain of SARS-CoV-2 (Corbett et al., 2020)].

For rapid characterization and monitoring of potential impact of virus, Centers for Disease Control and Prevention (CDC) has classified these variants into three categories, Variant of Interest (VOI), Variant of Concern (VOC) and Variant of High Consequences (VOHC) (https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Consequence). Variants were named after letters of the Greek alphabet by WHO and lineage and sub lineage of A, B, C, D by PANGO lineage system (Rambaut et al., 2020). Example of VOI is Epsilon (B.1.427, B.1.429), Eta (B.1.525), Iota (B.1.526), Kappa (B.1.617.1, B.1.617.3), Zeta (P.2) and Lambda (C.37). Depending on new scientific information, the status of a variant may escalate or deescalate. In United Kingdom (UK), B.1.1.7 variant with high transmission potential was identified as the first VOC 'Variant of Concern' and has said to be responsible for sudden increase in Covid-19 cases in UK in 2020 (Collier

* Corresponding author.
E-mail address: rokshana.parvin@bau.edu.bd (R. Parvin).

https://doi.org/10.1016/j.virol.2021.10.005
Received 12 August 2021; Received in revised form 2 October 2021; Accepted 18 October 2021
Available online 22 October 2021
0042-6822/© 2021 Elsevier Inc. All rights reserved.

et al., 2021). Later, this variant was designated according to CDC as Alpha variants with amino acid replacements in spike gene N501Y, P681H. Following this, two more VOC were identified in 2020; Beta (South African variant/B.1.351) with E484K, K417 N mutations and Gamma (Brazilian variant/P.1) with E484K mutations in spike gene (Gómez et al., 2021). Early in 2021, India was devastated with thousands of deaths daily and a new variant, named Delta/B.1.617.2, emerged as a highly contagious virus and was soon listed as a VOC in CDC classification of emerging variants. VOHC are variants with significantly low or no efficacy against available therapeutic or preventive therapy, but they have not yet been identified.

Within a year of the discovery of Covid-19 on March 8, 2020, Bangladesh was confronted with thousands of deaths in the first wave (https://iedcr.gov.bd/covid-19/covid-19-situation-updates). Bangladesh is in the midst of its second wave of pandemic, which began in March 2021. There is a sudden increase of detection rate as high as 30% per day in June–July 2021 with approximately 200 deaths (https://corona.gov.bd/graph). The disease pattern of the second wave has also shifted; most patients are hospitalized with breathing problems as well as very low oxygen saturation (SpO2), as opposed to the first wave, which was characterized by fever, throat pain, and coughing with moderately low SpO2. The high infectivity and transmission rate raised the topic of whether new SARS-CoV-2 genotypes could be introduced into Bangladesh. Furthermore, being a neighboring country of India, if the Delta variants spreads to Bangladesh, it would be worrying. Although, Bangladesh began its immunization campaign with the Covishield (Oxford-AstraZeneca) vaccine, which was followed by Sinopharm, Pfizer, and Moderna. So far, approximately 10% of the population has been fully vaccinated till August 2021 (https://www.unicef.org/bangladesh/en/press-releases/unicef-who-covax-ultra-low-temperature-freezers-boost-vaccine-storage-capacity-nine); however, a significant portion of the population remains unvaccinated, and reaching maximum (80%) coverage may take some time. The distribution and efficacy of the various vaccines, as well as their impact on VOC susceptibility in Bangladesh, are, however, unknown. Furthermore, the vaccine's efficacy against various VOCs is still being debated globally. Therefore, Bangladesh needs continuous monitoring to track the emergence of new variant as well as to identify the possible impact of those mutation on altered pathogenic potential. In our previous analytical study, we have reported the introduction of Alpha variants (UK; B.1.1.7) and dominant circulation of lineage B.1.1.25 in Bangladesh (Afrin et al., 2020; Parvin et al., 2021) during the year 2020. Genomic analysis is useful for tracing and tracking the distribution and evolution of circulatory viruses, which aids in the identification of new variants. Furthermore, novel variants such as VOC: Alpha, Beta, Gamma, and Delta (Tang et al., 2021; Tegally et al., 2021; Sabino et al., 2021; Singh et al., 2021a,b; O'Toole et al., 2021) as well as other variants, have been proven to impair the effectiveness of vaccination in convalescent or vaccinated individuals (Hoffmann et al., 2021; Davies et al., 2021; Sheikh et al., 2021). It is therefore critical to monitor VOCs and identify the local emerging variants through regular sequence analysis. Reporting of local variants is crucial in the campaign against Covid-19 and the spread of SARS-CoV-2. The aim of the study is to look at the dispersal of SARS-COV-2 variants circulating in Bangladesh at the first half of 2021, analyzing the prevalence of clades and lineages and determining the virus mutations.

2. Materials and methods

2.1. Retrieval of sequence metadata

The sequence data of SARS-CoV-2 genomes deposited in the Global Initiative on Sharing All Influenza Data (GISAID) EpiCoV (Shu and McCauley, 2017) database (https://www.gisaid.org/) was retrieved through filtering the time period from January to July 2021. Only complete genome sequences deposited from Bangladesh were selected and their corresponding FASTA, patient status metadata (Supplemental File 1) were downloaded from the GISAID platform. During the time period chosen, a total of 1325 partial and complete sequences were available, with only complete genome sequences [n = 820] obtained in the first half of 2021 being considered for analysis in this study.

2.2. Phylogenetic analysis

The file of sequence data in FASTA format was then imported into the web-based software Nextclade v1.5.2 (https://clades.nextstrain.org/) for quality assurance, clade assignment, lineage distribution, and diversity analysis. A phylogenetic tree was generated from 820 high quality Bangladeshi SARS-CoV-2 strains in comparison with the global reference strains (n = 1794) from different sub-continent that are automatically incorporated within the Nextclade platform (Hadfield et al., 2018). The generated data were further placed in Auspice v2.29.1 (https://auspice.us/) platform to check the phylogenetic diversity. Furthermore, the selected genomic sequences of Bangladeshi SARS-CoV-2 were aligned independently using multiple alignment (MAFFT v7) online platform (https://mafft.cbrc.jp/alignment/software/) (Katoh et al., 2019). A Randomized Axelerated Maximum Likelihood (RAxML) tree was built based on the general timereversible (GTR) model using RAxML v1.0.0 (Stamatakis, 2014) with 1000 bootstrap replicates and finally tree rendering was shown and furnished in iTOL (Letunic and Bork, 2019).

2.3. Descriptive analysis

Descriptive statistics was used to have the proportion and distribution of different variants/lineages/clades of SARS-CoV-2. A division-wise proportion of lineages map was generated to depict the virus distribution within the country using ArcGIS version 10.4. The graphs stating the month-wise proportion with 95% confidence interval were generated using an opensource web-based tool 'Datawrapper' (https://www.datawrapper.de).

2.3.1. Mutational profile

The 820 complete genome sequences of Bangladeshi SARS CoV-2 collected during the 1st half of 2021 were sequentially compared with the reference strain (hCoV-19/Wuhan/WIV04/2019) using CoVsurver algorithm available at the GISAID EpiCoV platform. The deduced amino acid (aa) substitutions at the spike (S) surface glycoprotein, in the large polyprotein 1 ab (nsp1-nsp16), four structural proteins (S, E, M, and N), and other accessory proteins (NS3, NS6, NS7a, NS7b, and NS8) were analyzed. The distribution and frequency of substitutions of the respective genes were examined carefully. The crystal structure of receptor binding domain S protein (aa position 336–518) from a Bangladeshi SARS CoV-2 was predicted on online based webtool Swiss-Model (https://swissmodel.expasy.org/interactive/58Ub4Q/models/) using best fit template from the repository (7krr.1.A). The projected model's PDB format was ornamented and the mutation spots were highlighted using UCSF *ChimeraX* software (Goddard et al., 2018).

3. Results

3.1. Demographic summary

According to patient metadata (n = 820) available from the database the nasopharyngeal and oropharyngeal samples were taken between January and June 2021. Males had a higher disease prevalence (66.1%) than females (33.9%), with a median age of 40 years for males and 38 years for females. The samples widely covered the 48 districts across 8 divisions of Bangladesh. Additional demographic and clinical data are summarized in Supplemental File 1.

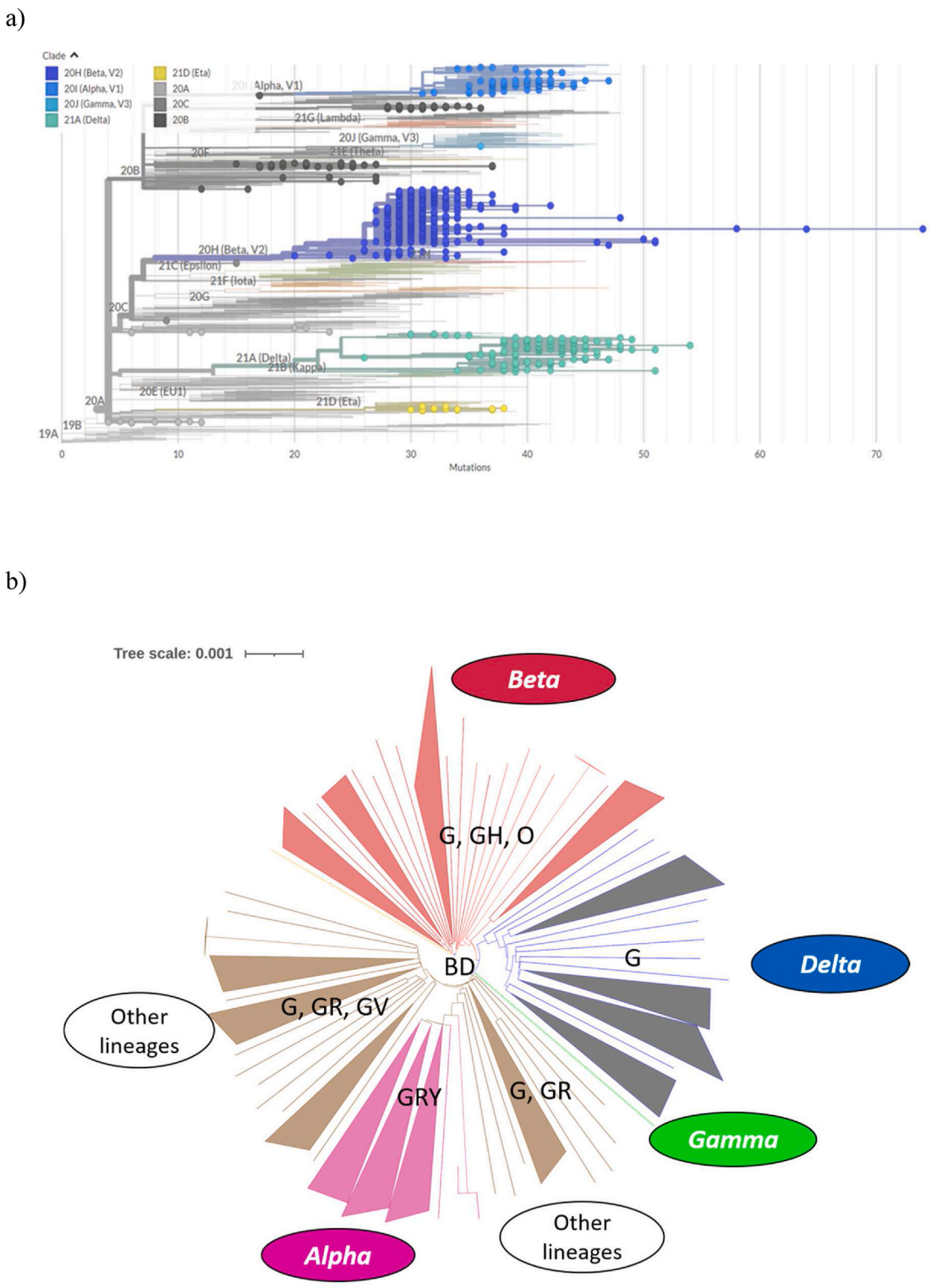


Fig. 1. Phylogenetic relationship of SARS-CoV-2 viruses from the ongoing novel coronavirus COVID-19 pandemic. a) Phylogenetic clusters and clades as generated by Nextstrain for the dataset of 820 high-quality Bangladeshi SARS CoV-2 genomes in comparison with the 1794 global reference sequences. The different color nodes (circles) indicated the designated clades (20H, 20I, 20 J, 21A, 21D, 20A, 20B and 20C) identified within the Bangladeshi strains. The global reference sequences are showing in ash color lines without any node circles. b) The circular representation of a phylogenetic tree constructed using the ML method with 820 Bangladeshi strains (BD) revealed separate clusters for different VOCs and lineages circulating in the country (Alpha: pink; Beta: red; Gamma: green; Delta: blue, and Others: brown). GISAID clades (G, GH, GR, GRY, GV, O) distribution were also depicted within each cluster. Taxa with similar node matrices within each branch were collapsed.

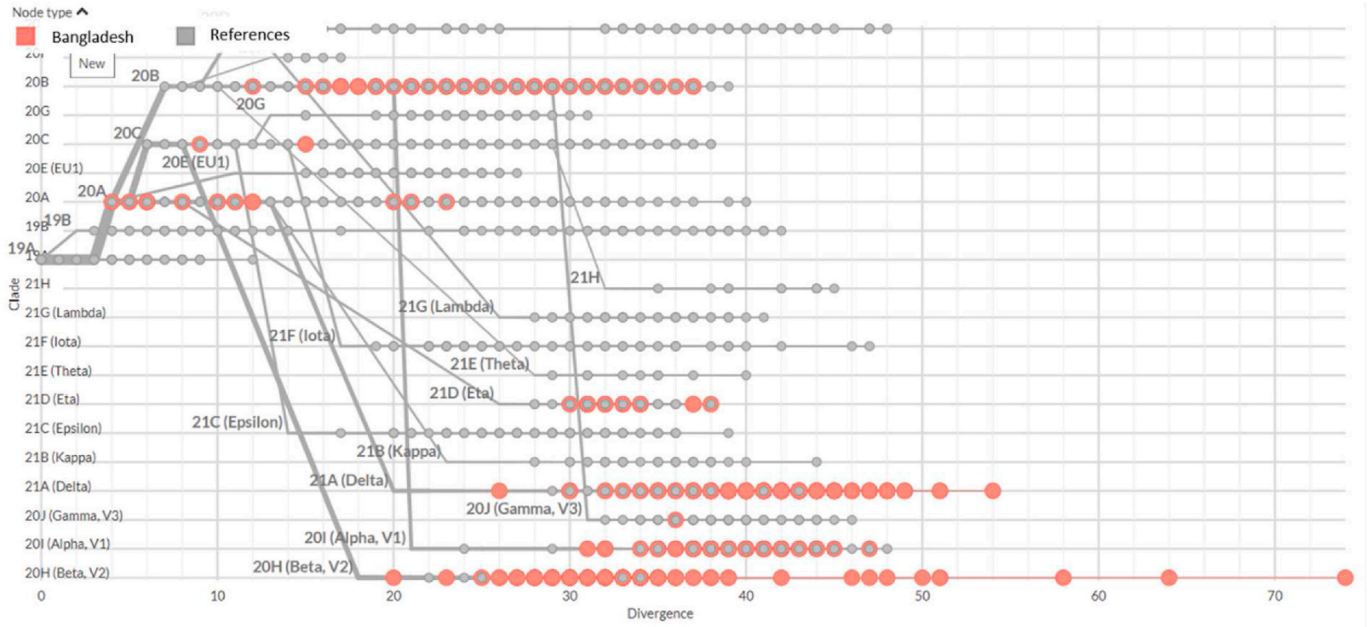


Fig. 2. Scatter plot showing evolutionary relationship and divergence limit based on Nextstrain phylogeny. The plot was generated in online Auspice platform (https://auspice.us/) by given input from Nextclade generated "JSON_Auspice" data. The branches showed the prototype strain from where the divergence start and pillared on the root strain (clade 19A) found at December 2019 in Wuhan (Wuhan-Hu-1/NC_045512). Bangladeshi strains were indicated by a red node. Until June 2021, the highest divergence was observed in clade 20H (Beta, V2).

3.2. Phylogenetic classification and diversity

Within the specified timeframe, genome sequences from Bangladesh revealed the presence of eight Nextstrain clades: 20A, 20B, 20C, 20H (Beta, V2), 20I (Alpha, V1), 20 J (Gamma, V3), 21A (Delta), 21D (Eta), and six GISAID clades: G, GH, GR, GRY, GV, and O (Fig. 1). Sequences further confirmed introduction of all four variants of concern (VOC) with a predominance of variants Alpha, Beta and Delta (Fig. 1b). VOC Gamma was identified in a single strain during the study period, indicating that community transmission had not yet occurred in the country with that particular VOC. In correlation with the PANGO lineages (Rambaut et al., 2020) it was observed that the lineage B.1.1.25, B.1.1.318, B.1.1.7/Alpha, B.1.351/Beta, B1. 617.2/Delta and B1.525/Eta were found to be widely distributed in the country. We further analyzed the diversity of SARS-CoV-2 within the variants and clades. Among the worldwide examined reference sequences, a Bangladeshi SARS-CoV-2 that belongs to 20H (Beta, V2) clade exhibits the highest diversity, strains from the 21A (Delta) and 20I (Alpha, V1) clades (Fig. 2) follow the 20H clade.

3.3. Distribution of Bangladeshi SARS CoV-2 variants, clades and lineages

The ongoing upsurge of COVID-19 pandemic in Bangladesh coincides with high prevalence of emerging variants, clades or lineages. At the beginning of 2021, GISAID clade GR and GRY were prevalent like previous year. Clade GH stands up in late February and takes over in March and April. In June, though, clade G overtakes clade GH and becomes more frequent (Fig. 3a). The observation was further supported by 95% confidence interval (CI) analysis of clade prevalence across the given time period. In the first two months of the year, the lineages B.1.1.25 and B.1.1.7/Alpha were prevalent, whereas B1.351/Beta became dominant in March and April. However, both introduced in the country in late 2020 (Table 1). The B.1.617.2/Delta variant surpassed Beta in May shortly after its inception in April, and it became widely visible in June (Fig. 3b). The months of February and May serve as a transitional period between three major variants (Alpha, Beta and Delta). In addition, another lineage B.1.318 along with the variant of interest (VOI) B.1.525/Eta that was detected in February and March respectively, remained steady at a low prevalence until June. The variant P.1/Gamma, on the other hand, was introduced in the country in February and has yet to be seen in later sequences. We further checked the spread of variants over the districts (second tier of regional administration) and divisions (first tier of regional administration) of Bangladesh. Dhaka, the capital of Bangladesh had the highest number of sequences in the GISAID database that identified almost all the lineages found in the country (Fig. 3c). The Delta variant that introduced in April, was found to be more common in the Rangpur, Rajshahi, Khulna, and Barishal divisions, which share a border with India.

3.4. Amino acid substitutions and relation to variants

A total of 3086 amino acid (aa) substitutions were observed in 820 Bangladeshi SARS-CoV-2 sequences (Supplemental File 2) analyzed in this study. Whereas most mutations were accumulating in respective lineages prior to January 2021, some have been repeated and have remained stable. The large polymerase gene (ORF 1 ab) that contains 16 proteins (nsp1-nsp16) had the most aa substitutions (n = 1853) particularly within nsp3 (n = 586), followed by the S protein (n = 541) and other proteins over the six months period as shown in Fig. 4a. The S protein receptor binding site (RBS) is a critical area in viral infection and pathogenesis, where a single mutation can result in the formation of a novel variant. All four VOCs; B.1.1.7/Alpha (484K, S494P, N501Y), B1.351/Beta (K417 N, E484K, N501Y), P.1/Gamma (K417T, E484K, N501Y), and B.1.617.2/Delta (K417 N, L452R, T478K), have been identified in Bangladeshi SARS-CoV-2, with K417 N, K417T, L452R, T478K, E484K, S494P, N501Y mutations at the RBS of S protein reported either as single, double, or triple mutant in the relevant strains. In Fig. 4b, the major mutations at RBS that result in emergence of VOC are represented in the structural model of S protein. In the current study, the VOI Eta (B.1.525 + E484K) as well as another unique variant B.1.1.318+ E484K were observed in the analyzed SARS-CoV-2 sequences. Although a newly emerging VOI, Lambda (C.37 lineage) that contains a mutation L452Q in the receptor-binding domain has not yet

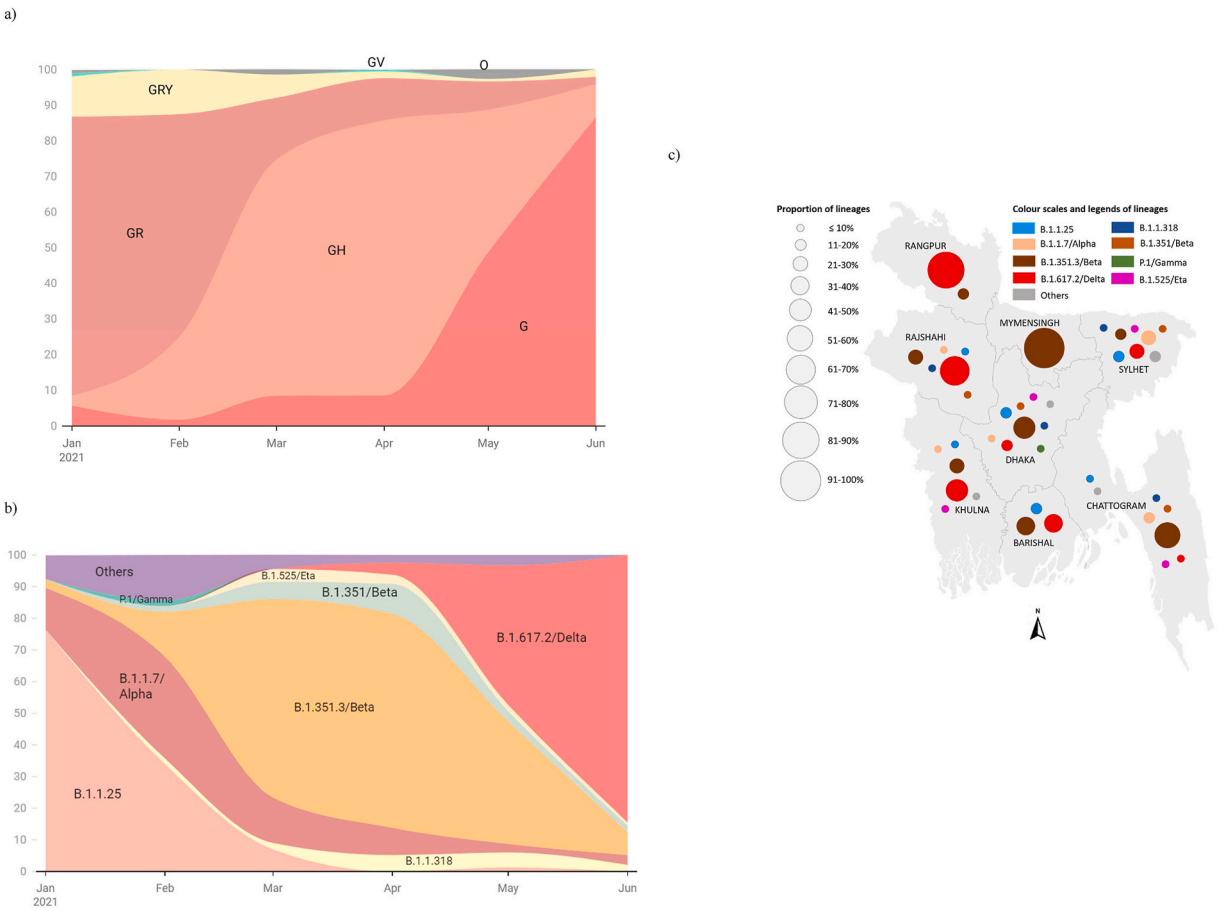


Fig. 3. Distribution of circulating clades and lineages and SARS CoV-2 surge within January to June 2021 in Bangladesh based on GISAID EpiCoV database. a) Stacked area plot shows the month-wise clade prevalence b) Stacked area plot shows the month-wise prevalence of various lineages c) The map showing the division-wise proportion of lineages within the given time frame. The proportion was estimated taking the total number of sequences accessible in each division as denominator.

been found in Bangladesh, the L452Q mutation was identified in a sequence in April in the present study. Table 1 lists the other major functional mutations found in addition to the RBS of S protein. The most important mutations observed in other structural and accessary proteins are also listed that observed frequently in global strains.

4. Discussion

Many efforts have been made to share information about the SARS CoV-2 through open platforms such as GISAID EpiCoV (Shu and McCauley, 2017) and Nextstrain (Hadfield et al., 2018) in order to better understand the molecular epidemiology, classification, and evolution of the virus in order to control and manage the disease. In this study, we looked at 820 full genome sequences of SARS-CoV-2 strains from the GISAID database, which were collected from various areas in Bangladesh between January and June 2021. However, the research is based on secondary source, it will contribute to a better understanding of the virus dissemination, prevalence of circulating variants, clades or lineages and summarize the mutations that have occurred in Bangladeshi SARS-CoV-2.

Our research observed a subsequent flip in the preponderance of variants in the first half of 2021 (January to June). According to the current sequence analysis, all the VOCs were introduced in Bangladesh (Fig. 1, Table 1). Alpha, Beta, and Delta VOCs switched quickly and overtook each other during this analysis period. Multiple introductions at the same time, as well as rapid community spread, may result in a quick flip. According to the Child Health Research Foundation (CHRF), 4/5-part sequences in February–March 2021 belonged to either the Alpha or Beta variant. (https://virological.org/t/detection-of-the-b-1-1-7-and-b-1-351-sars-cov-2-variants-in-bangladesh/668/1). In March 2021, the Alpha variant identification rate was 52% (https://www.icddrb.org/news-and-events/news?id=874) However, B1.351.3/Beta quickly took control, and it was discovered in 80% of all genome sequencing done by the International Centre for Diarrheal Disease Research in Bangladesh in the last two weeks of March. According to the Institute of Epidemiology, Disease Control and Research (IEDCR), Bangladesh, Beta variants were detected primarily from March to April, until when the development of Delta variants was confirmed in Bangladesh. Subsequently the Delta variant was detected in 45% sequences in May and 78% sequences in June in Bangladesh (https://old.iedcr.gov.bd/website/images/files/nCoV/FourthJulyCOVID19Update.pdf). All of these secondary data are in agreement with the results of our sequencing analysis. Recent news briefing from IEDCR, 98% of covid patient have been found to be affected with Delta variant despite the fact that a recent sequence is not yet available in the database. Circulation and rapid switching of new variants are thought to be responsible for potential transmission and re-infection even in vaccinated individuals (Singh et al., 2021a,b; Sabino et al., 2021; Hacisuleyman et al., 2021).

The current epidemic is even more severe than last year, with an official infection rate of around 30% (as of 31 July 2021) according to the daily update from Government (https://corona.gov.bd/). The reported cases and sequence analysis hint to the introduction of the lethal Delta variant by crossing the border (Fig. 4b), that wreaking havoc in India in early 2021 and now in Bangladesh. Dhaka, the capital of

Table 1
List of variants of concern and interest found in Bangladeshi SARS CoV-2 with their significant mutations and different lineages and clade designation.

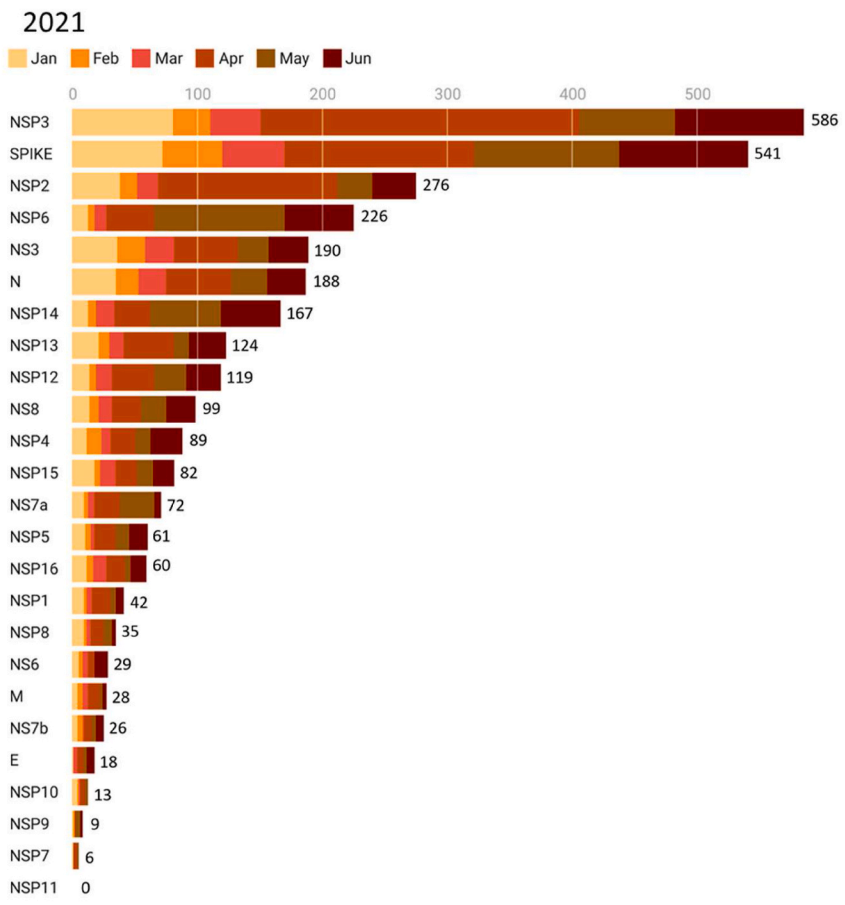
Variants and origin	Mutations in spike protein[a]	Mutations in other proteins[a]	PANGO Lineage	GISAID clade	Nexts-train clade	First detected in BD
VOC Alpha/UK variant	69del, 70del, 144del, **484K, S494P, N501Y,** A570D, D614G, P681H, T716I, S982A, D1118H, K1191 N	N: D3L, G204R, R203K, S235F, NS8: Q27stop, R52I, Y73C, NSP3: A890D, I1412T, T183I, NSP6: F108del, G107del, S106del, NSP12: P323L, NSP13:K460R, P78S	B.1.1.7	GRY	20I (V1)	31-12-20
VOC Beta/South African variant	D80A, D215G, 241del, 242del, 243del, **K417N, E484K, N501Y**, D614G, A701V	E: P71L, N: T205I, T362I, NS3: Q57H, S171L, NS7b: E39stop, NSP2: T85I, NSP3: K837 N, S794L, NSP5:K90R, NSP6: F108del, G107del, S106del	B.1.351 B.1.351.2 B.1.351.3	GH/ 501Y·V2	20H (V2)	1-11-20
VOC Gamma/ Brazilian variant	L18F, T20 N, P26S, D138Y, R190S, **K417T, E484K, N501Y,** D614G, H655Y, T1027I	N: G204R, P80R, R203K, NS3: S253P, NS8: E92K, NSP3: K977Q, P1103L, S370L, NSP4: S184 N, NSP6: F108del, G107del, S106del, NSP12: P323L, NSP13: E341D	P.1 P1.1 P1.2	GR/ 501Y·V3	20 J (V3)	20-04-21
VOC Delta/Indian variant	T19R, V70F, T95I, G142D, E156del, F157del, R158G, A222V, W258L, **K417N, L452R, T478K,** D614G, P681R, D950 N	N: D63G, D377Y, R203 M NS3: S26L NS7a: T120I, V82A NSP3: E545A, P822L NSP4: A446V, NSP6: T181I, V149A, NSP12: G671S, P323L, NSP13: P77L	B.1.617.2	G/ 478K·V1	21A	27-04-21
VOI Eta/Indian variant	A67V, 69del, 70del, 144del, **E484K,** D614G, Q677H, F888L	E: L21F M: I82T N: A12G, T205I NS3: S92L NSP3: E959A, K589 N, P1103S, T1189I NSP6: F108del, G107del, S106del, NSP12: P323F	B.1.525	G/ 484K·V3	21D	11.03.21

Bold indicates the mutation at receptor binding site that leads to emergence of variants.
Abbreviations; BD: Bangladesh, UK: United Kingdom,VOC: Variant of Concern, VOI: Variant of Interest.
[a] According to GISAID-EpiCOV mutation analysis. All mutation did not find in all strain.

Bangladesh and most densely populated city, has the most sequences in databases, indicating that it was the hub for the introduction and spread of all lineages and VOCs reported in the country. International airports and national gateways with numerous exit and entry points connected to the Indian border had modulated such a wide introduction and dispersion. Local trading and regular market visits for a livelihood may have accelerated the local and community transmission of VOCs in a short period of time.

Mutations in the SARS-CoV-2 spike (S) protein have received much attention in terms of monitoring emerging strains, because infection is driven by the S protein binding to human cell surface angiotensin converting enzyme 2 (ACE2) receptors (Conceicao et al., 2020). Receptor binding site (RBS) of S protein is the key element of virus-receptor interaction, host range and infectivity (Chen et al., 2020). Mutations at RBS in the SARS-CoV-2 genome that affect the S protein have emerged in a number of differentially-reported variants, including the E484K and N501Y mutations shared by the currently known VOCs, B.1.1.7/Alpha, B1.351.3/Beta, and P.1/Gamma. Another common RBS mutation is K417 N/T, which has been found in B1.351.3/Beta, P.1/Gamma, and B.1.617.2/Delta, whereas the B.1.617.2/Delta variant contains unique mutation L452R and T478K along with a mutation (K417 N) characteristic for other VOCs (Allen et al., 2021). The significant L452R mutation of B.1.617.2/Delta variant is associated with increased transmissibility and reduction in neutralization by convalescent plasma and specific therapeutic antibodies (Deng et al., 2021). Furthermore, the discovery of the L452Q mutation (typical of VOI C.37/Lambda) in the study sequences indicates that this VOI may appear in Bangladesh any time soon. D614G was previously identified as the most prevalent mutation in Bangladeshi strains (Afrin et al., 2020; Parvin et al., 2021) and is still present in all currently circulating variants (Table 1). Yet another alteration P681 R/H, which is located directly adjacent to the furin cleavage site (FCS) and may influence S1/S2 cleavage, cell entry, and infectivity (Helmy et al., 2020), has been found to be mutated in many variants (Nagy et al., 2021). In the current study, all of these mutations were found in Bangladeshi SARS-CoV-2 sequences, indicating that variants were introduced repeatedly before being transmitted to the community. Although the distribution of variants, lineages, or clades shown here is based on sequences from the database, we believe that this information will help in understanding the rapidly changing dynamics of SARS-CoV-2 and will aid in the development of effective control strategies.

a)



b)

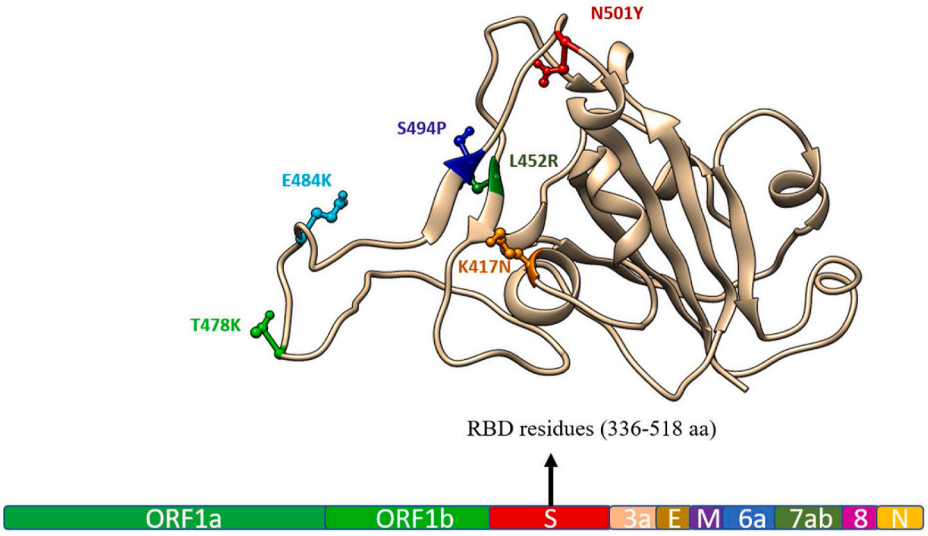


Fig. 4. Mutational profile of SARC-CoV-2 observed in Bangladesh from January to June 2021. a) Month-by-month distribution of amino acid alterations in the relevant proteins of Bangladeshi SARS-CoV-2. b) SARS-CoV-2 genomic structure with cartoon structural conformation of RBS (aa 336–518) of S protein. Differently colored level atoms indicated major mutations at the given position, which is also seen in Bangladeshi sequences and is required for VOC formulation.

5. Conclusions

In summary, this study investigated the phylodynamic, lineage, clade, or variant distribution, and prevalence of Bangladeshi SARS-CoV-2, as well as the mutation of the virus that was circulating in Bangladesh in January to July 2021. The prevalence of VOC B.1.1.7/Alpha, B1.351.3/Beta, and B.1.617.2/Delta and switching in two-month interval were the main highlight. However, by the end of June, only B.1.617.2/Delta was found in the majority of sequences, indicating that this is the main VOC to monitor in Bangladesh in the coming months.

CRediT authorship contribution statement

Sultana Zahura Afrin: wrote the first draft of the manuscript, organised data collection, designed the study, formal analysis, and interpreted the results, edited and critically reviewed the manuscript. **Md Taohidul Islam:** designed the study, formal analysis, and interpreted the results, edited and critically reviewed the manuscript. **Shyamal Kumar Paul:** edited and critically reviewed the manuscript, planned and co-supervision. **Nobumichi Kobayashi:** edited and critically reviewed the manuscript. **Rokshana Parvin:** wrote the first draft of the manuscript, organised data collection, designed the study, formal analysis, and interpreted the results, edited and critically reviewed the manuscript, planned and supervision of the entire work. All authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

The authors declare no potential conflict of interest.

Acknowledgements

The authors are thankful to the authority of Bangladesh Agricultural University Research System (BAURES), Mymensingh and Sheikh Hasina Medical College, Tangail, Bangladesh.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.virol.2021.10.005.

Author contributions

ASZ and RP wrote the first draft of the manuscript. RP and ASZ organised data collection, ASZ, MTI and RP designed the study, analyzed formally, and interpreted the results. ASZ, MTI, SKP, NK and RP edited and critically reviewed the manuscript. RP planned and supervised the work. All authors have read and agreed to the published version of the manuscript.

Financial support

This work did not receive any fund.

Ethics approval

Not Applicable.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Afrin, S.Z., Paul, S.K., Begum, J.A., Nasreen, S.A., Ahmed, S., Ahmad, F.U., Aziz, M.A., Parvin, R., Aung, M.S., Kobayashi, N., 2020. Extensive genetic diversity with novel mutations in spike glycoprotein of severe acute respiratory syndrome coronavirus 2, Bangladesh in late 2020. New Microbes New Infect 41, 100889. https://doi.org/10.1016/j.nmni.2021.100889.

Allen, H., Vusirikala, A., Flannagan, J., Twohig, K.A., Zaidi, A., Consortium, C.O.G.-U.K., Groves, N., Lopez-Bernal, J., Harris, R., Charlett, A., Dabrera, G., Kall, M., 2021. Increased Household Transmission of COVID-19 Cases Associated with SARS-CoV-2 Variant of Concern B. 1.617. 2: a National Case Control Study. Public Health England chrome-extension. http://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fkhub.net%2Fdocuments%2F135939561%2F405676950%2FIncreased%2BHousehold%2BTransmission%2Bof%2BCOVID-19%2BCases%2B-%2Bnational%2Bcase%2Bstudy.pdf%2F7f7764fb-ecb0-da31-77b3-b1a8ef7be9aa&clen=351317.

Chen, J., Wang, R., Wang, M., Wie, G.W., 2020. Mutations strengthened SARS-CoV-2 infectivity. J. Mol. Biol. 432 (19), 5212–5226. https://doi.org/10.1016/j.jmb.2020.07.009.

Chen, Y., Guo, Y., Pan, Y., Zhao, Z.J., 2020. Structure analysis of the receptor binding of 2019-nCoV. Biochem. Biophys. Res. Commun. 525 (1), 135–140. https://doi.org/10.1016/j.bbrc.2020.02.071.

Collier, D.A., De Marco, A., Ferreira, I., Meng, B., Datir, R., Walls, A.C., Kemp S, S.A., Bassi, J., Pinto, D., Fregni, C.S., Bianchi, S., Tortorici, M.A., Bowen, J., Culap, K., Jaconi, S., Cameroni, E., Snell, G., Pizzuto, M.S., Pellanda, A.F., Garzoni, C., Gupta, R.K., 2021. SARS-CoV-2 B.1.1.7 Sensitivity to mRNA Vaccine-Elicited, Convalescent and Monoclonal Antibodies. medRxiv: the Preprint Server for Health Sciences. https://doi.org/10.1101/2021.01.19.21249840.

Conceicao, C., Thakur, N., Human, S., Kelly, J.T., Logan, L., Bialy, D., Bhat, S., Stevenson-Leggett, P., Zagrajek, A.K., Hollinghurst, P., Varga, M., Tsirigoti, C., Tully, M., Chiu, C., Moffat, K., Silesian, A.P., Hammond, J.A., Maier, H.J., Bickerton, E., Shelton, H., Dietrich, I., Graham, S.C., Bailey, D., 2020. The SARS-CoV-2 Spike protein has a broad tropism for mammalian ACE2 proteins. PLoS Biol. 18 (12), e3001016 https://doi.org/10.1371/journal.pbio.3001016.

Corbett, K.S., Flynn, B., Foulds, K.E., Francica, J.R., Boyoglu-Barnum, S., Werner, A.P., Flach, B., O'Connell, S., Bock, K.W., Minai, M., 2020. Evaluation of the mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. N. Engl. J. Med. 383, 1544–1555.

Davies, N.G., Abbott, S., Barnard, R.C., Jarvis, C.I., Kucharski, A.J., Munday, J.D., Pearson, C., Russell, T.W., Tully, D.C., Washburne, A.D., Wenseleers, T., Gimma, A., Waites, W., Wong, K., van Zandvoort, K., Silverman, J.D., 2021. CMMID COVID-19 working group, COVID-19 genomics UK (COG-UK) consortium. Diaz-Ordaz, K., Keogh, R., … Edmunds, W. J. Science 372 (6538), eabg3055 https://doi.org/10.1126/science.abg3055. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England.

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., Reddy, N.P., Sanchez San Martin, C., Federman, S., Cheng, J., Balcerek, J., Taylor, J., Streithorst, J.A., Miller, S., Sreekumar, B., Chen, P.Y., Schulze-Gahmen, U., Taha, T. Y., Hayashi, J.M., Simoneau, C.R., Kumar, G.R., McMahon, S., Lidsky, P.V., Xiao, Y., Hemarajata, P., Green, N.M., Espinosa, A., Kath, C., Haw, M., Bell, J., Hacker, J.K., Hanson, C., Wadford, D.A., Anaya, C., Ferguson, D., Frankino, P.A., Shivram, H., Lareau, L.F., Wyman, S.K., Ott, M., Andino, R., Chiu, C.Y., 2021. Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. Cell 184 (13), 3426–3437. https://doi.org/10.1016/j.cell.2021.04.025 e8.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., Ferrin, T.E., 2018. UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci. 27 (1), 14–25. https://doi.org/10.1002/pro.3235.

Gómez, C.E., Perdiguero, B., Esteban, M., 2021. Emerging SARS-CoV-2 variants and impact in global vaccination programs against SARS-CoV-2/COVID-19. Vaccines 9 (3), 243. https://doi.org/10.3390/vaccines9030243.

Hacisuleyman, E., Hale, C., Saito, Y., Blachere, N.E., Bergh, M., Conlon, E.G., Schaefer-Babajew, D.J., DaSilva, J., Muecksch, F., Gaebler, C., Lifton, R., Nussenzweig, M.C., Hatziioannou, T., Bieniasz, P.D., Darnell, R.B., 2021. Vaccine breakthrough infections with SARS-CoV-2 variants. N. Engl. J. Med. 384 (23), 2212–2218. https://doi.org/10.1056/NEJMoa2105000.

Hadfield, J., Megill, C., Bell, S.M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T., Neher, R.A., 2018. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics 34, 4121–4123. https://doi.org/10.1093/bioinformatics/bty407.

Helmy, Y.A., Fawzy, M., Elaswad, A., Sobieh, A., Kenney, S.P., Shehata, A.A., 2020. The COVID-19 pandemic: a comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment, and control. J. Clin. Med. 9 (4), 1225 https://doi.org/10.3390/jcm9041225.

Hoffmann, M., Arora, P., Groß, R., Seidel, A., Hörnich, B.F., Hahn, A.S., Krüger, N., Graichen, L., Hofmann-Winkler, H., Kempf, A., Winkler, M.S., Schulz, S., Jäck, H.M., Jahrsdörfer, B., Schrezenmeier, H., Müller, M., Kleger, A., Münch, J., Pöhlmann, S., 2021. SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing antibodies. Cell 184 (9), 2384–2393. https://doi.org/10.1016/j.cell.2021.03.036 e12.

Katoh, K., Rozewicki, J., Yamada, K.D., 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings Bioinf. 20, 1160–1166.

Kumar, R., Harilal, S., Al-Sehemi, A.G., Pannipara, M., Behl, T., Mathew, G.E., Mathew, B., 2021. COVID-19 and domestic animals: exploring the species barrier crossing, zoonotic and reverse zoonotic transmission of SARS-CoV-2. Curr. Pharmaceut. Des. 27 (9), 1194–1201. https://doi.org/10.2174/1381612826666201118112203.

Letunic, I., Bork, P., 2019. Interactive Tree of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 47 (W1), W256–W259. https://doi.org/10.1093/nar/gkz239.

Li, T., Liu, D., Yang, Y., Guo, J., Feng, Y., Zhang, X., Cheng, S., Feng, J., 2020. Phylogenetic supertree reveals detailed evolution of SARS-CoV-2. Sci. Rep. 10 (1), 22366 https://doi.org/10.1038/s41598-020-79484-8.

Li, X., Wang, W., Zhao, X., Zai, J., Zhao, Q., Li, Y., Chaillon, A., 2020. Transmission dynamics and evolutionary history of 2019-nCoV. Transmission dynamics and evolutionary history of 2019-nCoV. J Med Virol. 2020. 92(5):501-511. doi:10.1002/jmv.25701 J Med Virol. May 92 (5), 501–511. https://doi.org/10.1002/jmv.25701.

Nagy, A., Basiouni, S., Parvin, R., Hafez, H.M., Shehata, A.A., 2021. Evolutionary insights into the furin cleavage sites of SARS-CoV-2 variants from humans and animals. Arch. Virol. 166 (9), 2541–2549. https://doi.org/10.1007/s00705-021-05166-z.

O'Toole, Á., Hill, V., Pybus, O.G., Watts, A., Bogoch, I.I., Khan, K., Messina, J.P., 2021. COVID-19 genomics UK (COG-UK) consortium, network for genomic surveillance in South Africa (NGS-SA), Brazil-UK CADDE genomic network. Tegally, H., Lessells, R. R., Giandhari, J., Pillay, S., Tumedi, K. A., Nyepetsi, G., Kebabonye, M., Matsheka, M., Mine, M., Tokajian, S., … Kraemer, M., 2021. Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 Wellcome open res 6, 121. https://doi.org/10.12688/wellcomeopenres.16661.1.

Parvin, R., Afrin, S.Z., Begum, J.A., Ahmed, S., Nooruzzaman, M., Chowdhury, E.H., Pohlmann, A., Paul, S.K., 2021. Molecular analysis of SARS-CoV-2 circulating in Bangladesh during 2020 revealed lineage diversity and potential mutations. Microorganisms 9 (5), 1035.

Rambaut, A., Holmes, E.C., O'Toole, Á., Hill, V., McCrone, J.T., Ruis, C., du Plessis, L., Pybus, O.G., 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nature Microbiol 5, 1403–1407. https://doi.org/10.1038/s41564-020-0770-5.

Sabino, E.C., Buss, L.F., Carvalho, M., Prete Jr., C.A., Crispim, M., Fraiji, N.A., Pereira, R., Parag, K.V., da Silva Peixoto, P., Kraemer, M., Oikawa, M.K., Salomon, T., Cucunuba, Z.M., Castro, M.C., de Souza Santos, A.A., Nascimento, V.H., Pereira, H. S., Ferguson, N.M., Pybus, O.G., Kucharski, A., Faria, N.R., 2021. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. Lancet 397 (10273), 452–455. https://doi.org/10.1016/S0140-.

Sheikh, A., McMenamin, J., Taylor, B., Robertson, C., Public Health Scotland and the EAVE II Collaborators, 2021. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. Lancet 397 (10293), 2461–2462. https://doi.org/10.1016/S0140-6736(21)01358-1.

Shu, Y., McCauley, J., 2017. GISAID: global initiative on sharing all influenza data - from vision to reality. Euro Surveill. 22, 30494 https://doi.org/10.2807/1560-7917.es.2017.22.13.30494.

Singh, J., Rahman, S.A., Ehtesham, N.Z., Hira, S., Hasnain, S.E., 2021a. SARS-CoV-2 variants of concern are emerging in India. Nat. Med. 27 (7), 1131–1133. https://doi.org/10.1038/s41591-021-01397-4.

Singh, J., Samal, J., Kumar, V., Sharma, J., Agrawal, U., Ehtesham, N.Z., Sundar, D., Rahman, S.A., Hira, S., Hasnain, S.E., 2021b. Structure-function analyses of new SARS-CoV-2 variants B.1.1.7, B.1.351 and B.1.1.28.1: clinical, diagnostic, therapeutic and public Health implications. Viruses 13 (3), 439. https://doi.org/10.3390/v13030439.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Tang, J.W., Tambyah, P.A., Hui, D.S., 2021. Emergence of a new SARS-CoV-2 variant in the UK. J. Infect. 82 (4), e27–e28. https://doi.org/10.1016/j.jinf.2020.12.024.

Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., Doolabh, D., Pillay, S., San, E.J., Msomi, N., Mlisana, K., von Gottberg, A., Walaza, S., Allam, M., Ismail, A., Mohale, T., Glass, A.J., Engelbrecht, S., Van Zyl, G., Preiser, W., de Oliveira, T., 2021. Detection of a SARS-CoV-2 variant of concern in South Africa. Nature 592 (7854), 438–443. https://doi.org/10.1038/s41586-021-03402-9.

Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y., Yuan, M.L., 2020. A new coronavirus associated with human respiratory disease in China. Nature 579 (7798), 265–269.

Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273.