GMR
The Original

GENOMIC SURVEILLANCE OF SARS-CoV-2: WHOLE-GENOME CHARACTERIZATION OF AN Alpha variant

Muhammad Umer[1], Muhammad Shafee[1,*], Manal Hadi Ghaffoori Kanaan[2], Sura Saad Abdullah[3], Zafar Ahmad[1], Muhammad Naeem[1], Mourad Ben Said[4,5], Farhad Badshah[6,7,*]

[1]Center for Advanced Studies in Vaccinology & Biotechnology (CASVAB), University of Balochistan, Quetta, Pakistan.
[2]Department of Food Industry Technologies/Technical Institute of Suwaria, Middle Technical University, Baghdad, Iraq.
[3]College of Pharmacy, Ajou University, Suwon 16499, Republic of Korea.
[4]Department of Basic Sciences, Higher Institute of Biotechnology of Sidi Thabet, University of Manouba, Manouba 2010, Tunisia.
[5]Laboratory of Microbiology, National School of Veterinary Medicine of Sidi Thabet, University of Manouba, Manouba 2010, Tunisia.
[6]Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Key Laboratory of Livestock and Poultry Multi-Omics of MARA, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518000, China.
[7]State Key Laboratory of Animal Biotech Breeding, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing 100193, China.
* Corresponding Author Muhammad Shafee (shafi.casvab@um.uob.edu.pk), Farhad Badshah (farhadbadshah999@gmail.com)

ABSTRACT

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genome sequencing remains essential for monitoring viral evolution, disease progression, control strategies, vaccine design, and therapeutic development. Balochistan, the largest province of Pakistan, shares borders with Iran and Afghanistan and experiences frequent cross-border movement, creating a strong need to characterize circulating regional variants. In this study, a nasopharyngeal swab from an adult male patient visiting the Public Health Laboratory, Quetta, in November 2021 was processed for whole-genome sequencing using the Illumina NextSeq500 next-generation sequencing platform. The isolate was assigned to the Alpha lineage and deposited under accession number OQ983908.1. Comparative analysis against the reference genome identified 20 mutations. Ten spike-protein mutations were detected, including three deletions and seven missense changes: H69, H70, Y144, A570D, N501Y, D614G, P681H, T716I, S982A, and D1118H. Six missense mutations occurred in ORF1ab, including N460Y, T1001I, A1708D, K1763N, I2230T, and P314L. Two mutations were identified in the nucleocapsid protein, G204R and S235F, and two occurred in ORF8, Q27- and Y73C. Phylogenetic analysis clustered the isolate with sequences from the USA, Italy, and Austria. These findings expand the epidemiological and evolutionary record of SARS-CoV-2 in Pakistan.

KEYWORDS: Alpha variant, Next-generation sequencing, SARS-CoV-2, Spike protein,

1. INTRODUCTION

Since its emergence in December 2019 in Wuhan, China, SARS-CoV-2 has undergone ongoing mutation, leading to the emergence of highly infectious genetic variants. These variants have significantly impacted the dynamics of the SARS-CoV-2 pandemic by increasing transmissibility, developing drug resistance, and facilitating viral adaptation (Pachetti et al., 2020). The epidemiology of COVID-19 can be better understood by examining the genomic changes of SARS-CoV-2 strains. Extensive efforts have been made by scientists to track the variations in the virus genome, which has played a significant role in shaping this pandemic. These studies have provided valuable insights into the viral genetic diversity and the transmission routes across different countries (Hadfield et al., 2018). Notably, multiple variants and subvariants of SARS-CoV-2, including Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), Omicron (B.1.1.529), have emerged, leading to millions of deaths and various waves of the pandemic (Saberiyan et al., 2022).

The SARS-CoV-2 genome has a length of approximately 29.9 kb and consists of various coding regions. The ORF1a/b region encodes sixteen non-structural proteins (NSP1 to NSP16) while the genome also includes four structural proteins comprising spike protein (S), matrix protein (M), envelope protein (E), and nucleocapsid phosphoprotein (N) (Al-Qaaneh et al., 2021; Yao et al., 2020). The spike protein plays a crucial role in transmission, pathogenicity, and evolution, making it highly relevant for vaccine development. In addition, RNA-dependent RNA polymerase (RdRp; nsp12), one of the key non-structural proteins encoded by ORF1ab, is responsible for viral RNA replication and transcription. Together with nsp7 and nsp8, it forms the core replication machinery, making it an important target for antiviral drug development (Abdullahi et al., 2020).

The first case of COVID-19 in Pakistan was detected on February 26th, 2020 (Jabeen et al., 2020). As of December 19, 2023, a total of 1,580,631 cases and 30,656 deaths have been confirmed in the country (Khurshid et al., 2023). Genomic epidemiological data from Pakistan has revealed the prevalence of different lineages during various waves of the pandemic. The B.1 lineage was dominant during the first wave, followed by the B.1.36 during the second wave. The B.1.1.7 (alpha) variant emerged as the most prevalent during the third wave of COVID-19 (Hassan et al., 2025). The Alpha variant was first reported in the United Kingdom in November 2020 and rapidly spread worldwide, including

Pakistan. By February 2021, it spread to 93 countries due to its high transmission rate, estimated from 50% to 70% (Kirby, 2021). In Pakistan, the Alpha variant emerged at the end of December 2020 and remained dominant until July 2021. During this period, the country experienced a significant increase in cases, with 335,728 cases and 7,849 deaths attributed to the variant (Fiaz et al., 2022). The Alpha variant accounted for 72.7% of the genomic incidence in Pakistan (Basheer & Zahoor, 2021). The spread of the Alpha variant in Europe and the UK triggered a second wave of the pandemic during the winter months of 2020. This resulted in millions of new cases and led to the imposition of a second lockdown in January 2021 (Fokas & Kastis, 2021; Grint et al., 2021).

Compared to ancestral viruses with the D614G mutation, the Alpha variant has accumulated 23 mutations. Among these, 14 are missense mutations distributed across ORF1ab (A1708D, I2230T, and T1001I), the S protein (A570D, N501Y, P681H, S982A, T716I and D1118H), ORF8 (Q27stop, Y73C and R52I), and the N protein (S235F and D3L). Additionally, there are six silent mutations (C913T, C5986T, C14676T, C15279T, and T16176C) found in ORF1ab and the M gene (T26801C). Furthermore, 3 deletion mutations are present: SGF 3675-3677del in ORF1ab and H69-V70del and Y144del in the S protein (Andrew, 2020). These extensive mutational changes in the Alpha variant, mainly in the spike protein, have been previously associated with various effects. A notable example of the impact of mutational changes in the S protein is the D614G mutation, which was first identified in mid-2020 and rapidly spread globally. By January 2021, over 95% of sequenced SARS-CoV-2 genomes contained this mutation. Today, the D614G mutation is present in all major SARS-CoV-2 strains and is linked to a substantial increase in the virus's infectivity (Chen et al., 2022; Korber et al., 2020).

Among the mutations in the S protein, N501Y is of particular significance as it occurs in the receptor-binding domain (RBD) that directly interacts with the human ACE2 (angiotensin-converting enzyme 2) receptor (Barnes et al., 2020). This mutation, located at spike position 501, one of six key contact residues in the RBD, can enhance binding affinity to hACE2, potentially increasing virus transmissibility (Leung et al., 2021; Zhao et al., 2021). Additionally, the S1/S2 furin cleavage site, not present in closely related coronaviruses, facilitates entry into respiratory epithelial cells and enhances transmission in animal models (Peacock et al., 2020; Zhu et al., 2020).

Several reports indicate that spike protein mutations can reduce susceptibility to neutralizing antibodies in clinical isolates (Andreano et al., 2021) and various *in-vitro* studies (Harvey et al., 2021). Additionally, deletions in amino acids Y144, Y145, and V146 in the NTD region of the S1 subunit of the spike protein can disrupt interactions with endogenous mAbs and affect cell entry (Dawood et al., 2021; Li et al., 2021).

These mutations likely have functional impacts due to their presence in highly conserved regions of the S protein. For instance, the B.1.1.7 lineage includes mutations such as C5388A (orf1ab: A1708D) and T24506G (Spike: S982A), which occur in conserved regions and may be associated with increased infectivity (Eslami et al., 2022; Jungreis et al., 2021). The A570D mutation in the spike SD1 region may modulate the opening and closing of the RBD (Yang et al., 2021). Another mutation, D1118H, found in the B.1.1.7 lineage, is associated with higher infectivity (McCarthy et al., 2021). The T716 mutation in the S1/S2 region has been linked to increased transmissibility and potential immune evasion (Lazarevic et al., 2021).

Outside the S protein, mutations in ORF1ab have been notable in the B.1.1.7 lineage. Studies have shown that the A1708D and I2230T mutations in ORF1ab can reduce CD8+ T cell activation and potentially contribute to immune evasion (Xiao et al., 2022). The T1001I mutation in ORF1a was also observed in genomes sequenced from India, indicating high mutation frequency. These findings suggest that while some mutations drive viral evolution, others may enhance viral adaptation in a specific population.

Mutation patterns in SARS-CoV-2 genomes may reflect factors such as population age distribution, gender, host immunity, and socioeconomic status (Beenish et al., 2023; Hu et al., 2024; Islam et al., 2023; Xiao et al., 2022; Xu et al., 2025; Zhang et al., 2020). The P314L mutation in ORF1ab, located near the drug-binding region in the hydrophobic cleft of RdRp, which is the main target of some antiviral drugs like remdesivir and favipiravir (Fiaz et al., 2022; Kumar et al., 2020). Occurrences of highly prevalent mutations in RdRp suggest that some therapeutically resistant viral strains are likely to emerge (Farkas et al., 2021).

The ORF8 Q27stop mutation in the B.1.1.7 lineage truncates or inactivates the ORF8 protein, potentially allowing downstream mutations to accumulate. Early in the pandemic, several isolates with deletions leading to loss of ORF8 expression were identified globally, including a significant cluster in Singapore with a 382 nt deletion that resulted in both truncated Orf7b and ablated ORF8 expression. The Singaporean strain associated with milder clinical infections and reduced post-infection inflammation eventually faded by the end of March following successful control measures in Singapore (Basheer & Zahoor, 2021). Subsequent studies have shown that the ORF8 deletion only modestly affects virus replication in human primary airway cells compared to viruses without the deletion, leading to a slight replication lag (Kirby, 2021). The stop codon at position 27 in the ORF8 protein of the B.1.1.7 lineage likely results in a loss of function. Additionally, the Y73C mutation in ORF8 has been linked to increased infectivity and transmissibility (Farkas et al., 2021; Li et al., 2021).

Among structural proteins, the nucleocapsid (N) protein is notably more stable and conserved. Mutations in the N protein of the B.1.1.7 variant include G204R and S235F. The G204R mutation, also present in the P.1 variant (also known as the Gamma variant), was first identified in the UK and Japan/Brazil. Along with R203K, G204R is one of the most frequently observed mutations in the N protein and is associated with increased N protein and sub-genomic RNA expression (Leary et al., 2020). The S235F mutation, located in the LKR region, alters epitope specificity, potentially impacting antibody responses and vaccine-induced protection (Mohammad et al., 2021).

Balochistan is the largest province, with almost 44% of its landscape located in the southern part of Pakistan. It shares two international borders with Afghanistan and Iran. There is a substantial monthly influx and outflow of tens of thousands of people crossing the border for trade, in addition to the visit of thousands of pilgrims to Iran during

holidays. Most of them return through the same route within a stipulated time, turning the unhealthy environment to spread infection through their cross-border movements especially during pandemic periods. Similarly, a large number of people also migrate to the Punjab and Sindh provinces for various reasons that may also favour viral transportation and circulation. To assess genomic diversity and identify the emergence of new and potentially dangerous viral strains, it is crucial to conduct surveillance of SARS-CoV-2.
This study aimed to evaluate the circulating strain in the subject population with relation to its mutational and genetic diversity.

2. MATERIALS AND METHODS

2.1. Ethics approval

The ethical approval committee of the University of Balochistan, Pakistan, ethically approved this research study (Approval No. UOB-CASVAB-2021-Ph.D-034). This study adhered to the Declaration of Helsinki. Informed verbal consent was obtained from the patient involved in the study. The consent process was witnessed by Mr. Saifur Rahman, a lab worker at the Provincial COVID Reference Lab, Quetta, and the verbal consent was documented in the patient's file. For confidentiality, all data were anonymized during the analysis. No minors were involved in this study.

2.2. Sample collection

The nasopharyngeal swab was obtained from an adult male suspected of having COVID-19 and was quarantined in quarantine center, Quetta, Pakistan. The collected sample was transported in viral transport medium (VTM, Capricorn) and sent to Provincial Public Health Laboratory (PPHL) of Fatima Jinnah Institute of Chest diseases, Quetta, Balochistan, Pakistan following international protocols for biosafety guidelines. While this study is based on a single isolate, it provides critical early genomic surveillance data from Balochistan, a region with limited sequencing resources. The region's high cross-border traffic makes it essential to understand circulating strains.

2.3. Viral RNA extraction

Total RNA was extracted from the specimen using NATCH CS2 (Sansure Biotech Inc, China) fully automated nucleic acid extraction system. This system is based on advanced magnetic bead technology (MB) and is designed to simplify the RNA extraction process. It utilizes robotic pipetting and magnetic beads to isolate high-quality RNA from a 500 µl sample. The NATCH CS2 system also incorporates an integrated UV lamp to minimize the risk of sample cross-contamination and employs an advanced heating and cooling system to maintain optimal temperature throughout the extraction process.

2.4. Viral RNA amplification and detection

The extracted RNA sample was subjected to viral amplification following the protocol provided by Sansure Biotech, China. The SARS-CoV-2 Multiplex Nucleic Acid Diagnostic Kit (Sansure Biotech Inc, China) was utilized for the detection of the COVID-19 virus. The kit employs multiplex real-time PCR technology to simultaneously detect three conserved target genes of the virus (ORF1ab, N, and E genes) along with an internal control gene. This approach enhances the accuracy and sensitivity of the test, reducing the risk of false negatives or positives. Following the preparation of the master mix, the RNA template, primers, nucleotides, and polymerase enzyme were added in the reaction tube according to the manufacturer's protocol. The reaction mixture was then distributed into PCR tubes and placed in a thermal cycler machine (Bio-Rad Laboratories, Inc., USA, CFX real-time PCR detection systems). RNA extraction and RT-qPCR were performed using the Sansure Novel Coronavirus (2019-nCoV) Nucleic Acid Diagnostic Kit according to the manufacturer's instructions. Quality control measures included the use of internal extraction controls, positive and negative controls. Amplification of the internal control confirmed successful RNA extraction and absence of PCR inhibitors. Results were interpreted based on Ct values in accordance with the manufacturer's recommended criteria.

2.5. Genome sequencing

Whole genome sequencing of the isolate from the indigenous clinical sample was carried out at a commercial facility using the next-generation sequencing (NGS) platform (Illumina NextSeq500) with the KAPA stranded RNA-Seq library preparation kit. Libraries were prepared from SARS-CoV-2 RNA. The workflow included the generation of double-stranded cDNA using a mixture of random and poly (T) priming. Quality assessment of sequencing data was performed using Illumina Sequencing Analysis Viewer (SAV). Demultiplexing and conversion of BCL files to FASTQ format were conducted using Illumina CASAVA version 1.8.2 (San Diego, USA). The reads were then mapped to the corresponding virus genome using Bowtie2 version 2.1.0 (Developed at Johns Hopkins University, USA).
Mutational analysis and lineage identification of the sequence were performed using the online tool at https://clades.nextstrain.org with the Wuhan-Hu-1/2019 (MN908947) sequence. Subsequently, a few sequences of WHO-designated variants were retrieved from NCBI, along with highest homology results from NCBI Nucleotide BLAST. For genome assembly, PRICE was used, and subsequent phylogenetic and molecular analyses of the S gene for the WHO-designated variants were conducted using MEGA.

3. RESULTS

3.1. Patient demographic data

A 35-year-old adult male suspected of having COVID-19 and was held in quarantine center of district Quetta, Balochistan, Pakistan was recruited in this study. The patient exhibited flu-like signs such as dry cough, fever, bone pain, fatigue, and asphyxia and nasopharyngeal swab samples were processed for viral genome sequence and mutational analysis.

3.2. Mutational analysis in our SARS-CoV-2 isolate

Whole-genome sequencing revealed that the SARS-CoV-2 genome obtained in this study belonged to the Pango B.1.1.7 (alpha) lineage. The sequence has been submitted to GenBank under accession number OQ983908.1 (SARS-CoV-2/QTA/Human/2021) and data can be retrieved through web link https://www.ncbi.nlm.nih.gov/nuccore/OQ983908.1 (**Figs**. **1 and 2**).

The analysis of the study isolate with Wuhan reference sequence identified 20 nucleotide mutations. Amino acid residue analysis using BioEdit version 7.0 tools revealed deletion mutations in the spike glycoprotein: H69del, V70del, and Y144del (located at positions 21992–21994). Other characteristic substitution mutations observed in the S protein included N501Y, A570D, D614G, P681H, T716I, S982A, and D1118H (**Table 1; Figs. 3 and 4**).

The spike protein exhibited the highest enrichment of mutations, with 3 deletion mutations and 7 amino acid substitution/missense mutations**.** The deletion mutations (H69-, V70-, Y144-) were located in the spike N-terminal domain. Three missense mutations (N501Y, A570D, D614G) were found in the receptor-binding domain (RBD) of the spike protein, which are characteristic mutations of the B.1.1.7 lineage (**Figs. 3 and 4**). Furthermore, mutations P681H, T716I, S982A, and D1118H were found in the S1/S2, UH (upstream helix), HR1 (heptad repeat 1), and CD (connector domain) regions of the spike protein, respectively (**Table 1**).

The ORF1ab (Open Reading Frame 1ab) region exhibited a total of 6 missense mutations and no deletion mutations. These mutations included N460Y, T1001I, A1708D, K1763N, I2230T and P314L. The Nucleocapsid (N) and ORF8 protein had two mutations each. The N protein had two missense mutations (G204R and S235F), while ORF8 had one missense (Y73C) and one stop-gained (Q27-) mutation.

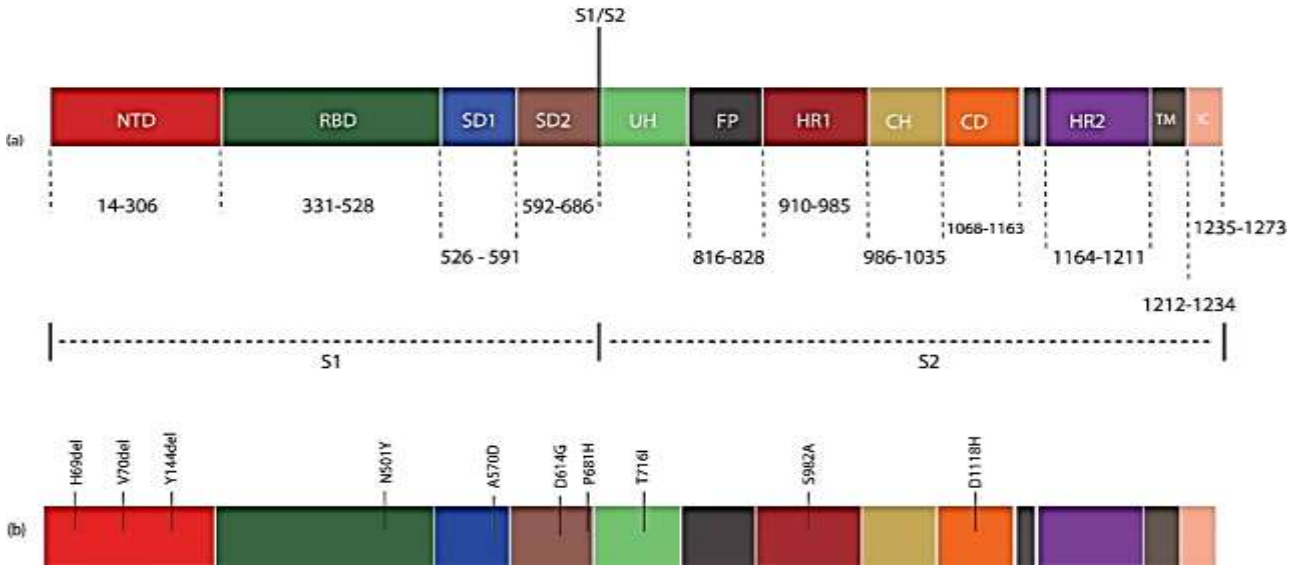


Fig. 1. The representative scheme of **(a)** Wuhan reference sequence (Accession ID: NC_045512). **(b)** Mutational profile in SARS-CoV-2/QTA/Human/2021 in comparison with Wuhan strain

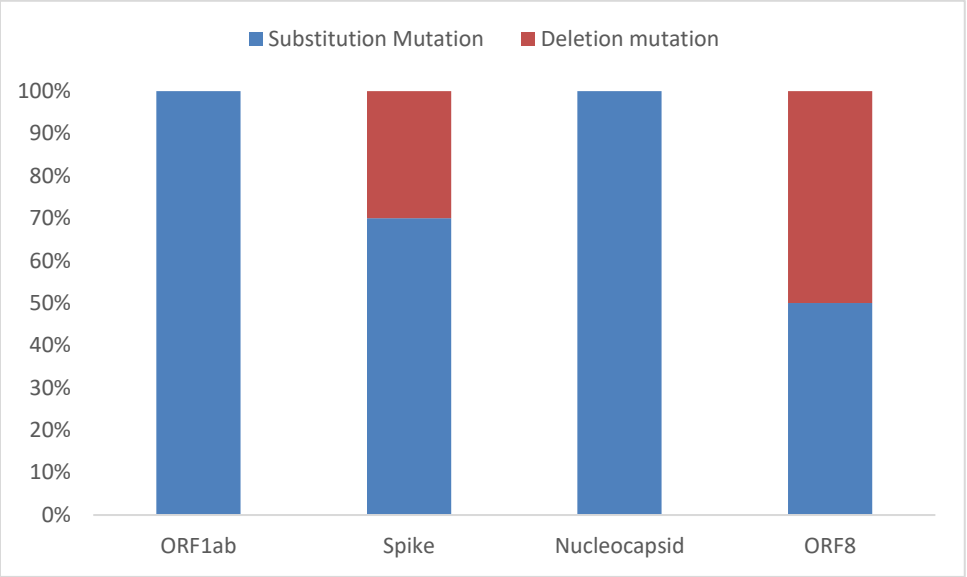


Fig. 2. Graphical presentation of mutational abundance detected in various proteins of local isolate (SARS-CoV-2/QTA/Human/2021).

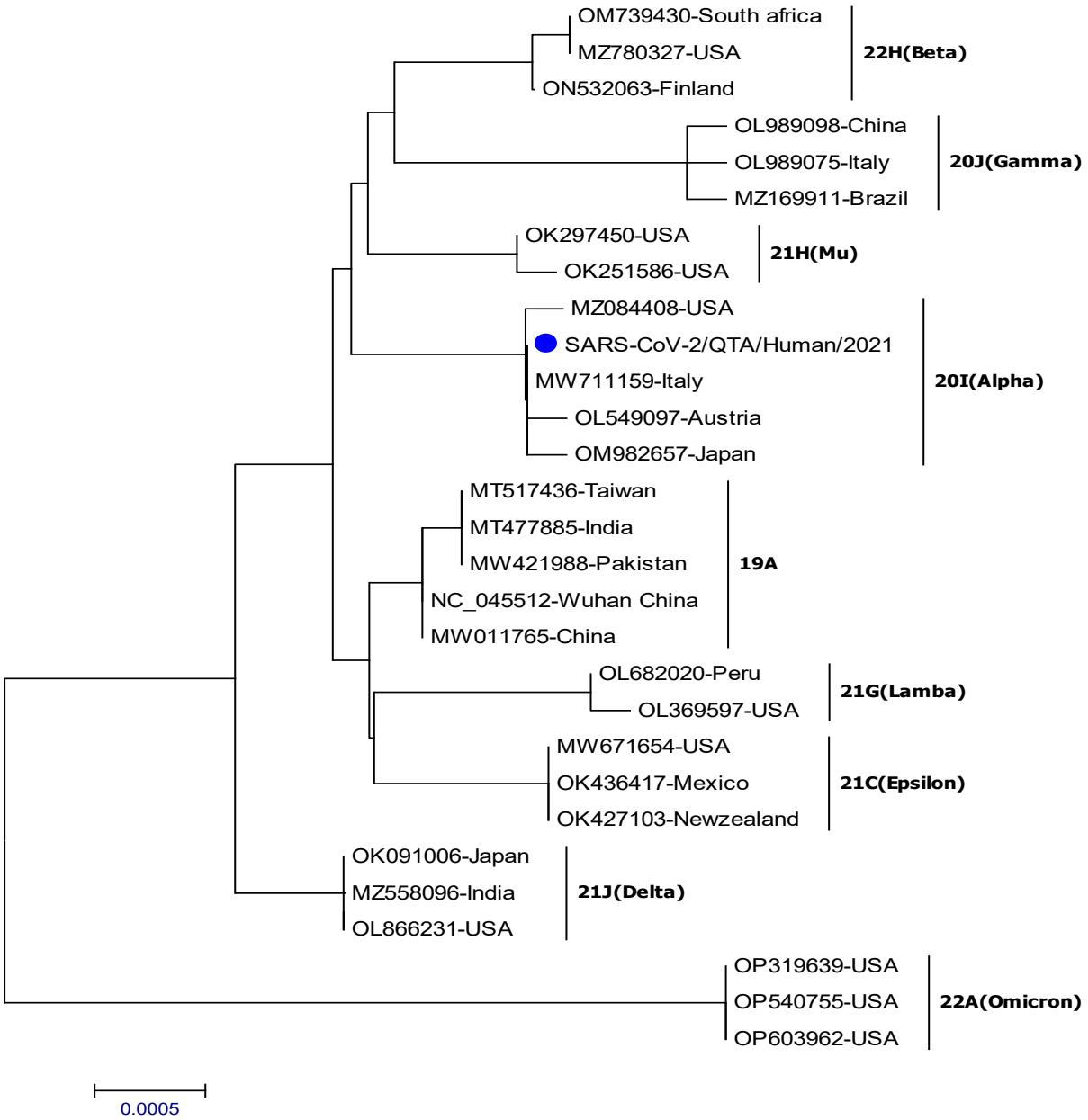


Fig. 3. Phylogenetic tree of our local isolate "SARS-CoV-2/QTA/Human/2021" based on Spike gene sequence

compared with global strains (BioEdit version 7.0).

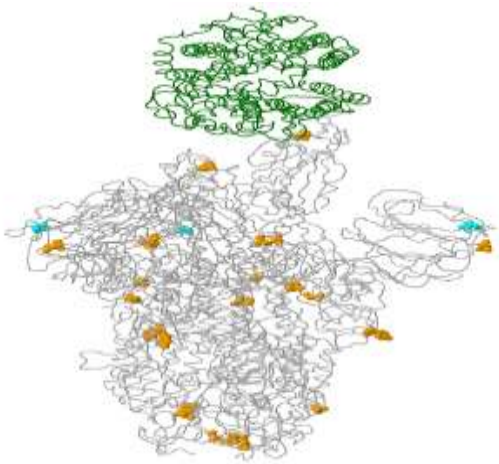


Fig. 4. Result for comparison with reference selection: Wuhan reference strain WIV04: 3D structural visualization of the spike glycoprotein with amino-acid changes identified in our sequences shown as colored balls. Spike glycoprotein (PDB: 6acj, EM 4.2 Angstrom) in complex with host cell receptor ACE2 (green ribbon). List of variations displayed in structure (nearest residue if in loop/termini region). **H69del V70del(69)Y144del(143)N501Y A570D D614G P681H(674) T716I S982A D1118H,** Using Spike glycoprotein mutation surveillance of GISAID website.

Table 1. Mutation profile in SARS-CoV-2/QTA/Human/2021 in comparison with Wuhan reference sequence (Accession ID: NC_045512).

Gene	Protein	Nucleotide Mutation	Amino Acid Mutation	Mutation Type	Possible functional implications
ORF1ab	-	A1643T	N460Y	Missense	Extremely rare mutation (200~ occurrences in GISAID). (Shapira et al., 2022). Functional implications are unclear
	-	C3267T	T1001I	Missense	Might be responsible virus evolution and adaptation (Islam et al., 2023).
	Nsp3	C5388A	A1708D	Missense	Might be associated with relatively high infectivity (Jungreis et al., 2021; Meng et al., 2021; Eslami et al., 2021). Decrease in CD8+ T cell activation and a possible immune evasion (Xiao et al., 2022).
	-	G5554T	K1763N	Missense	Functional implications are unclear
	-	T6954C	I2230T	Missense	Decrease in CD8+ T cell activation and a possible immune evasion (Xiao et al., 2022).
	RdRp	C14408T	P314L	Missense	Might contribute to antiviral drug resistance (Kumar et al., 2020; Koyama et al., 2020; Fiaz et al., 2022)
Spike	NTD	Gap; 21765-21770	Del HV69-70	Deletion	Enhances viral infectivity and could also contribute to antibody evasion (Meng et al., 2021)
	NTD	Gap;21992-21994	Del Y144	Deletion	Confers resistance to mAb (Suryadevara, 2021)
	RBD	A23063T	N501Y	Missense	Antibody escape (Harvey et al., 2021). Enhanced binding of SARS-CoV-2 spike protein to the human ACE2 receptor (Luan et al., 2021)
	RBD	C23271A	A570D	Missense	Might contribute to immune evasion of variants (Lazarevic, 2021)
	RBD	A23403G	D614G	Missense	Increases virion spike density and infectivity (Zhang et al., 2020). Increases viral replication in the upper respiratory tract and enhances the vulnerability of the virus to neutralizing antibodies (nAbs) (Plante, 2021; Zhang, 2020).
	S1/S2	C23604A	P681H	Missense	May increase spike cleavage by furin-like proteases, this does not significantly impact viral entry or cell–cell spread (Lubinski et al., 2021)
	S1/S2	C23709T	T716I	Missense	Might be associated with increased transmissibility and potential immune evasion (Lazarevic, 2021)
	HR1	T24506G	S982A	Missense	Associated with relatively high infectivity (Meng et al., 2021).

	CD	G24914C	D1118H	Missense	Associated with relatively high infectivity (Meng et al., 2021).
ORF8	ORF8	C27972T	Q27-	Stop-gained	Truncates and presumably inactivates ORF8 and causes further downstream mutations (Young et al. 2020; Gamage et al. 2020; Jungreis et al., 2021).
	ORF8	A28111G	Y73C	Missense	Increases infectivity and transmissibility of the virus (Farkas et al., 2021; Li et al., 2021).
N	N	G28883C	G204R	Missense	G204R with R203K is associated with high-transmissibility (Wu et al., 2021) and increase in N protein and sub-genomic RNA expression (Leary et al., 2021).
	N	C28977T	S235F	Missense	Associated with alteration of corresponding epitopes, affecting the specificity antibodies and vaccine-induced immunity (Mohammad et al., 2021).

ORF: Open reading frame**; S:** Spike glycoprotein**; N:** Nucleocapsid phosphoprotein**; RdRp:** RNA dependent RNA

3.3. Phylogenetic study and comparison to other variants

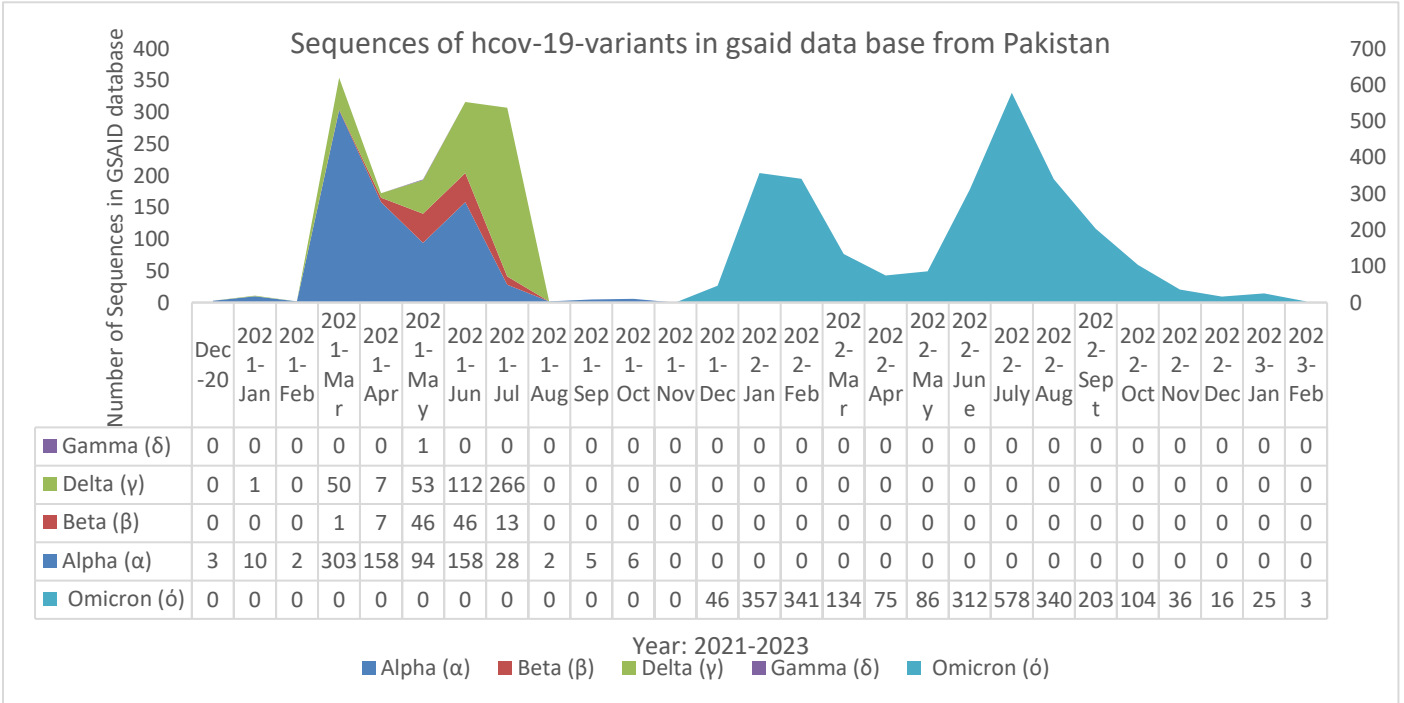
The complete genome sequence of the study isolate SARS-CoV-2/QTA/Human/2021, along with previously described variants of SARS-CoV-2, was aligned and compared to the Wuhan reference sequence (Accession ID: NC_045512) using Nextstrain tools (https://clades.nextstrain.org). The whole genome sequences of various SARS-CoV-2 variants used in this analysis were retrieved from NCBI GenBank. Phylogenetic analysis of the complete genome sequence revealed that the study isolate clustered with isolates from the UK (Accession ID: OW522582), USA (Accession ID: MZ084408), Italy (Accession ID: MW711159), Japan (Accession ID: OM982657), and Germany (Accession ID: OW998408) (**Fig. 3**).

The phylogeny analysis showed that our local SARS-CoV-2 genome belonged to clade 20I and was clustered closely with strains reported from the UK (Accession ID: OW522582), USA (Accession ID: MZ084408), Italy (Accession ID: MW711159), Japan (Accession ID: OM982657) and Germany (Accession ID: OW998408).

This phylogenetic tree (**Fig. 3**) indicates that the study isolate belongs to the B.1.1.7 (alpha) lineage. Sequences were retrieved from NCBI based on WHO-designated variants. To illustrate the epidemiology of the strain, sequences from the study isolate lineage were selected based on maximum homology from NCBI BLAST and used to construct the phylogenetic tree.

3.4. Distribution of variants of concern over time in Pakistan

The distribution pattern of different variants of concern (VOC) of SARS-CoV-2 over time in Pakistan after analyzing 4027 genomes submitted to GISAID database from Pakistan during December 2020 to February 2023 is shown in **Fig. 5**. Results clearly revealed that our sample was collected in November 2021 during the tail end of 3rd wave in Pakistan, and was followed by Omicron during 2022-2023.



	Dec-20	2021-Jan	2021-Feb	2021-Mar	2021-Apr	2021-May	2021-Jun	2021-Jul	2021-Aug	2021-Sep	2021-Oct	2021-Nov	2021-Dec	2022-Jan	2022-Feb	2022-Mar	2022-Apr	2022-May	2022-June	2022-July	2022-Aug	2022-Sept	2022-Oct	2022-Nov	2022-Dec	2023-Jan	2023-Feb
Gamma (δ)	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Delta (γ)	0	1	0	50	7	53	112	266	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta (β)	0	0	0	1	7	46	46	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Alpha (α)	3	10	2	303	158	94	158	28	2	5	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Omicron (ó)	0	0	0	0	0	0	0	0	0	0	0	0	46	357	341	134	75	86	312	578	340	203	104	36	16	25	3

Fig. 5. Graphical presentation of temporal dynamics of variant abundance over time in Pakistan clearly shows that our sample was collected from tail end of 3rd wave (Source: http:// gisaid.org/hcov-19- variants statistics. Accessed on _23.09.2024.Using excel spreadsheet).

4. DISCUSSION

In Pakistan, the first SARS-CoV-2 whole-genome sequence (SARS-CoV-2/ Gilgit1/human/2020/PAK) was submitted to GenBank on March 25, 2020 (Saif et al., 2021). In this study, we sequenced a SARS-CoV-2 genome from Quetta, Balochistan that belonged to B.1.1.7 (Alpha) lineage (clade 20I). It is plausible that Alpha was circulating in Quetta during the third wave but remained undocumented due to restricted sampling coverage, logistical constraints, or reliance on mutation-specific RT-PCR assays instead of comprehensive genome sequencing.

Variants such as Alpha (α), Beta (β), Gamma (γ), Delta (δ) and Omicron (O) are classified as variants of concern (VOC) due to their increased transmissibility and disease severity, resulting in a higher rate of hospitalizations (WHO). This Alpha variant was the primary driver of the second wave of the pandemic in the UK and other European countries, leading to a second lockdown in 2021-22. In Pakistan, it triggered in third wave, with a 97.9% incidence of the B.1.1.7 variant in Lahore, Pakistan. Moreover, whole-genome sequencing revealed that 87.5% (7/8) of Pakistani isolates were identical to variants from the UK, Switzerland, and other European countries (Sarwar et al., 2021).

Previous studies conducted in Pakistan have shown that the Alpha variant was predominant during the third wave of infections and was less abundant during the fourth wave. The genomic sequence of the Alpha variant was first submitted to the GISAID on December 25, 2020. During the third wave of COVID-19 in Pakistan, the B.1.1.7 variant was predominantly observed from February to June 2021, with the highest number of cases confirmed on April 17, 2021 (Sarwar et al., 2021; Umair et al., 2021). Similarly, the same variant was also reported in Turkey (Akçeşme et al., 2022) and in the United States (Bolze et al., 2022). The genomic sequences of the third wave in major cities of Pakistan primarily consisted of the B.1.1.7 variant, which accounted for 69% of the genomes. However, it is noteworthy that in this study, no Alpha variant was observed in Quetta, Balochistan, where our study was conducted (Shakeel et al., 2022).

In our current study, we investigated the mutational profile in the B.1.1.7 (Alpha) variant and identified a total of 20 mutations. The spike protein had the highest number of mutations, accounting for 50% of the total mutations, followed by the ORF1ab region, which accounted for 30% of the mutations. The ORF8 and nucleocapsid (N) proteins both exhibited a similar proportion of mutational abundance, with each accounting for approximately 10% of the mutations. Out of the 20 mutations, 16 were missense (non-synonymous) mutations, three were deletion mutations (two in the spike and one in ORF8), and one was a stop-gained mutation.

In this study, we identified several mutations in the Alpha variant (B.1.1.7.), including three deletion mutations (del. HV69-70 and Y144) and seven missense mutations (N501Y, D614G, A570D, T716I, P681H, S982A and D1118H) in the S protein. The RBD of the spike protein harboured three mutations (N501Y, A570D, and D614G), which are defining mutations of B.1.1.7 lineage. Consistent with our findings, previous studies by Umair et al. (Umair et al., 2022) and Aziz et al. (Aziz et al., 2023) reported characteristic mutations in the spike protein of the Alpha variant, including H69-, A570D, N501Y, D1118H, S982A, T716I, and P681H. It is noteworthy that SARS-CoV-2 has undergone numerous substitution mutations, especially in the spike protein. According to the GISAID database, the spike protein has more than 1229 amino acid substitutions. The characteristic mutations of the Alpha variant (B.1.1.7.), such as Δ144 (s), G204R (N), T1001I and 12230T (ORF1ab), were absent in the isolate from Lahore, Pakistan, despite showing ≥ 90% similarity to the globally prevalent variant. Notably, the P314L mutation, located near the drug-binding region in the cleft of RdRp, the target site for antiviral drugs, suggests the potential for the emergence of therapeutic-resistant viral strains (Dawood et al., 2021).Understanding these substitutions in spike amino acids across different lineages is crucial for assessing their potential impact on virulence, vaccine efficacy, and viral transmission. Therefore, continuous monitoring and surveillance of lineages and variants are essential for the development of effective public health strategies to control this pandemic (Guruprasad, 2022).

Particularly, the N501Y mutation has been reported to enhance viral binding to the hACE2 receptor, thereby increasing transmissibility (Galloway et al., 2021), and enabling the virus to evade the antibodies (Starr et al., 2020).With increased transmissibility by 43–90% (Davies et al., 2021) and approximately twice the replicative advantage (Grabowski et al., 2021), the Alpha variant quickly surpassed the Wuhan strain by spreading rapidly to many countries around the world.

In our local strain under study, we detected significant D614G mutation in the spike protein. Zhang et al. (2020) demonstrated this mutation is associated with enhanced viral loads worldwide. *In-vitro* studies conducted by Li et al. (2020) using pseudo viruses showed that the D614G mutation is correlated with enhanced infectivity. Similarly, among the spike variants, the G614 mutant exhibited the highest cell entry (Ozono et al., 2020).

We also identified the G204R and S235F mutations in the nucleocapsid (N) protein, which is crucial for RNA activities such as replication and plays a key role in RNA binding. Another mutation, **P314L,** was detected and is a target of some antiviral drugs such as favipiravir and remdesivir (Koyama et al., 2020; Kumar et al., 2020). Similarly, some characteristic mutations that we detected in our local strain, such as T1001I (ORF1ab), G204R (N), and Δ144(S), were not found in a set of 461 Alpha sequences in Pakistan (Fiaz et al., 2022).However, most of the mutations we detected in the current strain genome, including H69-, D614G, A570D, S982A, D1118H, T716I, and P681H in the spike protein; **P314L** in RdRp; S235F in N; Q27-, Y73C in ORF8, have been reported to have a prevalence of more than 95% globally. These findings highlight the presence of specific mutations in our local strain that may contribute to viral infectivity, replication, and potential interactions with antiviral drugs. It is essential to continue monitoring and analyzing the mutational landscape of SARS-CoV-2 to understand the implications for viral characteristics, transmission, and therapeutic interventions.

Fascinatingly, our sequence exhibited some unique and rare mutations of interest, including N460Y, T1001I, and

K1763N in the ORF1ab region. The N460Y mutation is particularly rare, with approximately 200 occurrences reported in the GISAID database. Its presence in 100% of viral reads suggests it may confer a significant fitness advantage, despite being difficult to explain through selective pressure (Shapira et al., 2022)**.** However, little is known about these mutations, which warrants further investigation.

The distribution pattern of variants shows that the Alpha variant was predominantly present towards the end of the third COVID wave in Pakistan, followed closely by the Omicron variant, which emerged as a variant of concern, as illustrated in Fig 5. Clinically, most of the patients in the province remained asymptomatic, whereas a few presented with fever, body ache, headache and cough. The mortality rate in the Province was low as compared to other provinces of Pakistan. This may be attributed to limited testing facilities and poor access to health care settings in the province, as supported by Badini et al. (Badini et al., 2021). The phylogenetic analysis of our current isolate showed clustering with sequences originating from various countries such as the UK (Accession ID: OW522582), USA (Accession ID: MZ084408), Italy (Accession ID: MW711159), Japan (Accession ID: OM982657), and Germany (Accession ID: OW998408). This suggests that the viral lineage in our study has genetic similarities with strains circulating in these countries. Furthermore, it indicates that the Alpha variant, which was primarily transmitted from different countries including the United Kingdom, has also reached Balochistan. In a study conducted by Umair et al**. (2021)** in Pakistan, two Alpha variant sequences were analyzed, and they exhibited 99.9% nucleotide similarity with the UK strain (VOC-202012/01). The phylogenetic analysis revealed two separate lineages, one clustering with genomes from Spain and the other with genomes from the USA. Similarly, a study from Nigeria by Olorunfemi et al. (Olorunfemi et al., 2024) reported the circulation of four different variants, in clade 20I, with the Alpha variant being the most predominant (33%). The current study, along with similar studies conducted in Pakistan, provides valuable insights into the genetic diversity, evolution, and transmission patterns of SARS-CoV-2 in the country. Understanding the local strain's sequence information is crucial for acquisition of better insights into viral epidemiology, immunogenicity, and pathogenicity. Overall, the study contributes valuable evidence to fill an epidemiological gap in Balochistan and supports the need for sustained sequencing efforts to better understand viral evolution and transmission dynamics at the provincial level.

Limitation

The primary limitation of this study was the limited number of samples analyzed within the specified geographical area. However, the findings offer valuable baseline data for ongoing surveillance of SARS-CoV-2 in the region.

CONCLUSION

This study provides a valuable baseline for future research by exploring the molecular epidemiology and genomic diversity of the B.1.1.7. variant. Our findings offer insights into its mutational profile of a single isolate circulating in Quetta, Balochistan, Pakistan, with the spike protein showing the highest number of mutations, followed by ORF1ab. Notably, we identified key mutations, including D614G in the spike protein, which is linked to increased viral loads and enhanced infectivity. The presence of these characteristic Spike mutations suggests a high prevalence of this variant, underscoring the need for regular monitoring to ensure preparedness for future pandemics in the region. Phylogenetic analysis revealed clustering with sequences from various countries, indicating global transmission dynamics. Our findings, along with similar studies in Pakistan, contribute to a better understanding of genetic diversity, evolution, and transmission patterns of SARS-CoV-2. These findings are crucial for comprehending viral epidemiology, immunogenicity, and pathogenicity, emphasizing the importance of continuous surveillance and sequence analysis of local strains for future pandemic preparedness.

DECLARATION

Authors' contributions

Conceptualization, M.S. and Z.A.; methodology, M.U; software, M.U., F.B., M.H.G.K., S.S.A. and M.B.S; validation, M.S., M.U., M.N., Z.A., M.B.S. and F.B.; formal analysis, M.U., F.B., M.S., M.H.G.K, and S.S.A.; investigation, M.U.; resources, F.B., Z.A. and M.S; data curation, M.U., M.N., and Z.A.; writing—original draft preparation, M.U., F.B., M.B.S. and M.S.; writing—review and editing, M.S., M.U., M.N., M.H.G.K, S.S.A, Z.A., MBS. and F.B; visualization, M.U., and F.B; supervision, M.S. and Z.A.; project administration, M.S. and F.B; funding acquisition,

F.B. All authors have read and agreed to the final version of the manuscript.

Funding

This study did not receive funding.

Acknowledgement

The authors would like to express their gratitude for the support and guidance provided by the lab in-charge at Fatima Jinnah TB Hospital, Quetta, and Prof. Dr. Tahir Yaqoob, in charge of the BSL3 Lab at UVAS, Lahore, Pakistan, and Dr. Zubair Shabbir, Professor at UVAS, for their invaluable technical assistance.

Consent for publication

All authors have read and reviewed the manuscript and agreed to its publication**.**

Data Availability Statement

The datasets generated and/or analyzed during the current study are available in the GenBank repository, accession number OQ983908.1, at https://www.ncbi.nlm.nih.gov/nuccore/OQ983908.1.

Conflict of interest
All authors declared that they have no conflict of interest.

REFERENCES

1. Abdullahi IN, Emeribe AU, Ajayi OA, Oderinde BS, et al. (2020). Implications of SARS-CoV-2 genetic diversity and mutations on pathogenicity of the COVID-19 and biomedical interventions. J. Taibah Univ. Med. Sci. 15: 258-264. https://doi.org/10.1016/j.jtumed.2020.06.005

2. Akçeşme FB, Köprülü TK, Erkal B, İş Ş, et al. (2022). Tracking the circulating SARS-CoV-2 variants in Turkey: complete genome sequencing and molecular characterization of 1000 SARS-CoV-2 samples. bioRxiv. https://doi.org/10.1101/2022.04.19.488722

3. Al-Qaaneh AM, Alshammari T, Aldahhan R, Aldossary H, et al. (2021). Genome composition and genetic characterization of SARS-CoV-2. Saudi J. Biol. Sci. 28: 1978-1989. https://doi.org/10.1016/j.sjbs.2020.12.053

4. Andreano E, Piccini G, Licastro D, Casalino L, et al. (2021). SARS-CoV-2 escape from a highly neutralizing COVID-19 convalescent plasma. Proc. Natl. Acad. Sci. U. S. A. 118: e2103154118. https://doi.org/10.1073/pnas.2103154118

5. Andrew R (2020). Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. Virological.

6. Aziz MW, Mukhtar N, Anjum AA, Mushtaq MH, et al. (2023). Genomic diversity and evolution of SARS-CoV-2 lineages in Pakistan. Viruses. 15: 1450. https://doi.org/10.3390/v15071450

7. Badini AM, Badini A, Mengal NM and Nanji K (2021). Characteristics of patients presenting with COVID-19 from Balochistan Province and lessons learnt. J. Coll. Physicians Surg. Pak. 31: S104-S108.

8. https://doi.org/10.29271/jcpsp.2021.Supp2.S104

9. Barnes CO, Jette CA, Abernathy ME, Dam KA, et al. (2020). SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. Nature. 588: 682-687. https://doi.org/10.1038/s41586-020-2852-1

10. Basheer A and Zahoor I (2021). Genomic epidemiology of SARS-CoV-2 divulge B.1, B.1.36, and B.1.1.7 as the most dominant lineages in first, second, and third wave of SARS-CoV-2 infections in Pakistan. Microorganisms. 9:

11. 2609. https://doi.org/10.3390/microorganisms9122609

12. Beenish K, Saira F, Zahida P, Mohammad A, et al. (2023). COVID-19 and SARS-CoV-2: what we know so far. J. Popul. Ther. Clin. Pharmacol. 30: 1745-1760. https://doi.org/10.53555/jptcp.v30i17.2854

13. Bolze A, Luo S, White S, Cirulli ET, et al. (2022). SARS-CoV-2 variant Delta rapidly displaced variant Alpha in the United States and led to higher viral loads. Cell Rep. Med. 3: 100564. https://doi.org/10.1016/j.xcrm.2022.100564

14. Chen Z, Li J, Li T, Fan T, et al. (2022). A CRISPR/Cas12a-empowered surface plasmon resonance platform for rapid and specific diagnosis of the Omicron variant of SARS-CoV-2. Natl. Sci. Rev. 9: nwac104.

15. https://doi.org/10.1126/science.abg3055

16. Davies NG, Abbott S, Barnard RC, Jarvis CI, et al. (2021). Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science. 372: eabg3055. https://doi.org/10.1126/science.abg3055

17. Dawood RM, El-Meguid MA, Salum GM, El-Wakeel K, et al. (2021). Bioinformatics prediction of B and T cell epitopes within the spike and nucleocapsid proteins of SARS-CoV-2. J. Infect. Public Health. 14: 169-178.

18. https://doi.org/10.1016/j.jiph.2020.12.006

19. Eslami S, Glassy MC and Ghafouri-Fard S (2022). A comprehensive overview of identified mutations in SARS CoV-2 spike glycoprotein among Iranian patients. Gene. 813: 146113. https://doi.org/10.1016/j.gene.2021.146113

20. Farkas C, Mella A, Turgeon M and Haigh JJ (2021). A novel SARS-CoV-2 viral sequence bioinformatic pipeline has found genetic evidence that the viral 3′ untranslated region is evolving and generating increased viral diversity. Front. Microbiol. 12: 665041. https://doi.org/10.3389/fmicb.2021.665041

21. Fiaz N, Zahoor I, Saima S and Basheer A (2022). Genomic landscape of alpha-variant of SARS-CoV-2 circulated in Pakistan. PLoS One. 17: e0276171. https://doi.org/10.1371/journal.pone.0276171

22. Fokas AS and Kastis GA (2021). SARS-CoV-2: the second wave in Europe. J. Med. Internet Res. 23: e22431. https://doi.org/10.2196/22431

23. Galloway SE, Paul P, MacCannell DR, Johansson MA, et al. (2021). Emergence of SARS-CoV-2 B.1.1.7 lineage

- United States, December 29, 2020-January 12, 2021. MMWR Morb. Mortal. Wkly. Rep. 70: 95-99.

24. https://doi.org/10.15585/mmwr.mm7003e2

25. Grabowski F, Preibisch G, Giziński S, Kochańczyk M, et al. (2021). SARS-CoV-2 variant of concern 202012/01 has about twofold replicative advantage and acquires concerning mutations. Viruses. 13: 392.

26. https://doi.org/10.3390/v13030392

27. Grint DJ, Wing K, Williamson E, McDonald HI, et al. (2021). Case fatality risk of the SARS-CoV-2 variant of concern B.1.1.7 in England, 16 November to 5 February. Euro Surveill. 26: 2100256. https://doi.org/10.2807/1560-7917.Es.2021.26.11.2100256

28. Guruprasad K (2022). Mutations in human SARS-CoV-2 spike proteins, potential drug binding and epitope sites for COVID-19 therapeutics development. Curr. Res. Struct. Biol. 4: 41-50. https://doi.org/10.1016/j.crstbi.2022.01.002

29. Hadfield J, Megill C, Bell SM, Huddleston J, et al. (2018). Nextstrain: real-time tracking of pathogen evolution. Bioinformatics. 34: 4121-4123. https://doi.org/10.1093/bioinformatics/bty407

30. Harvey WT, Carabelli AM, Jackson B, Gupta RK, et al. (2021). SARS-CoV-2 variants, spike mutations and immune escape. Nat. Rev. Microbiol. 19: 409-424. https://doi.org/10.1038/s41579-021-00573-0

31. Hassan ZU, Park M, Park D, Amin H, et al. (2025). Nanopore sequencing reveals the genomic diversity of the variants of concern of SARS-CoV-2 during 2021 disease outbreak in Pakistan. Sci. Rep. 15: 28129.

32. https://doi.org/10.1038/s41598-025-12774-1

33. Hu YL, Zhang YJ, Lv XY, Liu RL, et al. (2024). Impact of omicron variant infection on female fertility and laboratory outcomes: a self-controlled study. Am. J. Reprod. Immunol. 92: e70012.

34. Islam MA, Marzan AA, Arman MS, Shahi S, et al. (2023). Some common deleterious mutations are shared in SARS-CoV-2 genomes from deceased COVID-19 patients across continents. Sci. Rep. 13: 18644.

35. https://doi.org/10.1038/s41598-023-45517-1

36. Jabeen K, Haider H, Haider Z, Ali S, et al. (2020). Coronavirus (COVID-19) pandemic: outbreak, current scenario, and impact on human physiology in Pakistan. https://doi.org/10.20944/preprints202009.0040.v1

37. Jungreis I, Sealfon R and Kellis M (2021). SARS-CoV-2 gene content and COVID-19 mutation impact by comparing 44 Sarbecovirus genomes. Nat. Commun. 12: 2642. https://doi.org/10.1038/s41467-021-22905-7

38. Khurshid B, Khan RNA, Ahmad A and Shah YM (2023). COVID-19 and SARS-CoV-2: what we know so far. J. Popul. Ther. Clin. Pharmacol. 30: 1745-1760. https://doi.org/10.53555/jptcp.v30i17.2854

39. Kirby T (2021). New variant of SARS-CoV-2 in UK causes surge of COVID-19. Lancet Respir. Med. 9: e20-e21. https://doi.org/10.1016/s2213-2600(21)00005-9

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

41. https://doi.org/10.1016/j.cell.2020.06.043

42. Koyama T, Platt D and Parida L (2020). Variant analysis of SARS-CoV-2 genomes. Bull. World Health Organ. 98: 495-504. https://doi.org/10.2471/blt.20.253591

43. Kumar BK, Venkatraja B, Prithvisagar KS, Rai P, et al. (2020). Mutational analysis unveils the temporal and spatial distribution of G614 genotype of SARS-CoV-2 in different Indian states and its association with case fatality rate of COVID-19. bioRxiv. Preprint. https://doi.org/10.1101/2020.07.27.222562

44. Lazarevic I, Pravica V, Miljanovic D and Cupic M (2021). Immune evasion of SARS-CoV-2 emerging variants: what have we learnt so far? Viruses. 13: 1192. https://doi.org/10.3390/v13071192

45. Leary S, Gaudieri S, Chopra A, Pakala S, et al. (2020). Three adjacent nucleotide changes spanning two residues in SARS-CoV-2 nucleoprotein: possible homologous recombination from the transcription-regulating sequence. bioRxiv. Preprint. https://doi.org/10.1101/2020.04.10.029454

46. Leung K, Shum MH, Leung GM, Lam TT, et al. (2021). Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020. Euro Surveill. 26: 2002106. https://doi.org/10.2807/1560-7917.Es.2020.26.1.2002106

47. Li Q, Wu J, Nie J, Zhang L, et al. (2020). The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. Cell. 182: 1284-1294.e1289. https://doi.org/10.1016/j.cell.2020.07.012

48. Li X, Zhang L, Chen S, Ji W, et al. (2021). Recent progress on the mutations of SARS-CoV-2 spike protein and suggestions for prevention and controlling of the pandemic. Infect. Genet. Evol. 93: 104971. https://doi.org/10.1016/j.meegid.2021.104971

49. McCarthy KR, Rennick LJ, Nambulli S, Robinson-McCarthy LR, et al. (2021). Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. Science. 371: 1139-1142. https://doi.org/10.1126/science.abf6950

50. Mohammad T, Choudhury A, Habib I, Asrani P, et al. (2021). Genomic variations in the structural proteins of SARS-CoV-2 and their deleterious impact on pathogenesis: a comparative genomics approach. Front. Cell. Infect. Microbiol. 11: 765039. https://doi.org/10.3389/fcimb.2021.765039

51. Olorunfemi AB, Suliman SAR, Tran TT, Ayorinde B, et al. (2024). Whole genome sequencing and phylogenetic analysis of SARS-CoV-2 strains isolated during the COVID-19 pandemic in Nigeria. IJID Reg. 10: 174-178. https://doi.org/10.1016/j.ijregi.2024.01.005

52. Ozono S, Zhang Y, Ode H, Tan TS, et al. (2020). Naturally mutated spike proteins of SARS-CoV-2 variants show differential levels of cell entry. bioRxiv. https://doi.org/10.1101/2020.06.15.151779

53. Pachetti M, Marini B, Benedetti F, Giudici F, et al. (2020). Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. J. Transl. Med. 18: 179. https://doi.org/10.1186/s12967-020-02344-6

54. Peacock TP, Goldhill DH, Zhou J, Baillon L, et al. (2020). The furin cleavage site of SARS-CoV-2 spike protein is a key determinant for transmission due to enhanced replication in airway cells. bioRxiv. Preprint. https://doi.org/10.1101/2020.09.30.318311

55. Saberiyan M, Karimi E, Khademi Z, Movahhed P, et al. (2022). SARS-CoV-2: phenotype, genotype, and characterization of different variants. Cell Mol. Biol. Lett. 27: 50. https://doi.org/10.1186/s11658-022-00352-6

56. Saif R, Mahmood T, Ejaz A, Zia S, et al. (2021). Whole genome comparison of Pakistani Corona virus with Chinese and US strains along with its predictive severity of COVID-19. Gene Rep. 23: 101139. https://doi.org/10.1016/j.genrep.2021.101139

57. Sarwar MB, Yasir M, Alikhan NF, Afzal N, et al. (2021). SARS-CoV-2 variants of concern dominate in Lahore, Pakistan in April 2021. Microb. Genom. 7: 000693. https://doi.org/10.1099/mgen.0.000693

58. Shakeel M, Irfan M, Nisa ZU, Farooq S, et al. (2022). Genome sequencing and analysis of genomic diversity in the locally transmitted SARS-CoV-2 in Pakistan. Transbound. Emerg. Dis. 69: e2418-e2430. https://doi.org/10.1111/tbed.14586

59. Shapira L, Lerner S, Assayag G, Vardi A, et al. (2022). Discovery of novel spike/ACE2 inhibitory macrocycles using in silico reinforcement learning. Front. Drug Discov. 2: 1085701. https://doi.org/10.3389/fddsv.2022.1085701

60. Starr TN, Greaney AJ, Hilton SK, Ellis D, et al. (2020). Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. Cell. 182: 1295-1310.e1220. https://doi.org/10.1016/j.cell.2020.08.012

61. Umair M, Ikram A, Rehman Z, Haider SA, et al. (2022). Genomic diversity of SARS-CoV-2 in Pakistan during the fourth wave of pandemic. J. Med. Virol. 94: 4869-4877. https://doi.org/10.1002/jmv.27957

62. Umair M, Ikram A, Salman M, Alam MM, et al. (2021). Importation of SARS-CoV-2 variant B.1.1.7 in Pakistan.

J. Med. Virol. 93: 2623-2625. https://doi.org/10.1002/jmv.26869

63. Xiao C, Mao L, Wang Z, Gao L, et al. (2022). SARS-CoV-2 variant B.1.1.7 caused HLA-A2(+) CD8(+) T cell epitope mutations for impaired cellular immune response. iScience. 25: 103934. https://doi.org/10.1016/j.isci.2022.103934

64. Xu Z, Li EH, Liu J, Zhang YJ, et al. (2025). Postpartum hemorrhage emerges as a key outcome of maternal SARS-CoV-2 omicron variant infection surge across pregnancy trimesters. J. Infect. Public Health. 18: 102733.

65. Yang TJ, Yu PY, Chang YC, Liang KH, et al. (2021). Effect of SARS-CoV-2 B.1.1.7 mutations on spike protein structure and function. Nat. Struct. Mol. Biol. 28: 731-739. https://doi.org/10.1038/s41594-021-00652-z

66. Yao H, Lu X, Chen Q, Xu K, et al. (2020). Patient-derived mutations impact pathogenicity of SARS-CoV-2. medRxiv. Preprint. https://doi.org/10.1101/2020.04.14.20060160

67. Zhang L, Jackson CB, Mou H, Ojha A, et al. (2020). The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. bioRxiv. Preprint. doi: 10.1101/2020.06.12.148726. https://doi.org/10.1101/2020.06.12.148726

68. Zhao S, Lou J, Cao L, Zheng H, et al. (2021). Quantifying the transmission advantage associated with N501Y substitution of SARS-CoV-2 in the UK: an early data-driven analysis. J. Travel Med. 28: taab011. https://doi.org/10.1093/jtm/taab011

69. Zhu Y, Feng F, Hu G, Wang Y, et al. (2020). The S1/S2 boundary of SARS-CoV-2 spike protein modulates cell entry pathways and transmission. bioRxiv. Preprint. doi: 10.1101/2020.08.25.266775. https://doi.org/https://doi.org/10.1101/2020.08.25.266775