How to Cite:
Ul Abedien, Z., Ur Rahman, S., Shafiq, M., Gul, K., Tanveer, U., Mehmood, A., Zafar, N., Ali, S., & Sajid, S. (2022). Analysis of receptor binding domain for possible mutations in S gene region of SARS-CoV-2. *International Journal of Health Sciences*, *6*(S9), 4267–4278.
https://doi.org/10.53730/ijhs.v6nS9.13688

Analysis of receptor binding domain for possible mutations in S gene region of SARS-CoV-2

Zain Ul Abedien
Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan

Sajjad Ur Rahman
Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan

Maheen Shafiq
Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan

Kainat Gul
Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan and Department of Microbiology, Abbottabad University of Science and Technology, Abbottabad, Pakistan

Ujalla Tanveer
Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan

Ahad Mehmood
Department of Microbiology, Abbottabad University of Science and Technology, Abbottabad, Pakistan

Nishat Zafar
Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan

Sultan Ali
Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan

Sanaullah Sajid
Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan
Corresponding author email: sanaullahsajid@gmail.com

***Abstract*---**Humanity has historically been affected by remarkable epidemics and pandemics, including the plague, cholera, influenza, SARS-CoV, and MERS-CoV. A novel coronavirus pandemic known as SARS-CoV-2 is rapidly sweeping the globe. Over the period, the genome of the novel coronavirus has been mutated as it passes

International Journal of Health Sciences ISSN 2550-6978 E-ISSN 2550-696X © 2022.
Manuscript submitted: 9 July 2022, Manuscript revised: 18 Sept 2022, Accepted for publication: 27 Oct 2022

through its primary host. The world is reporting multiple point mutations. So the objective of this study was to observe modifications in the partial region of SARS-CoV-2 for the surveillance. This cross sectional study was carried out to detect the modifications in the SARS-CoV-2. For this purpose initial screening of COVID-19 was done to collect the strain of SARS-CoV-2 using RT-PCR. Then, primers were created in order to amplify the area using the S gene of the SARS-CoV-2. The amplified product was sent for the sequencing and bioinformatic tools were used to observed the mutations and data was compared with the Wild strain of the Virus. During the analysis, one of the most important point mutation was D614G caused by the amino acid substitution of aspartic acid with the glycine. Addition to this point mutation, other important mutations have also been observed. In this research, E484k, N501Y and D614G mutations were seen in the partial receptor binding domain (RBD) and SD1 region of the S gene sequence. As a result of point mutations, the amino acids were altered leading to enhanced transmission, binding affinity to the human ACE2 receptor, and reduced susceptibility against the antibody neutralization. Globally, these alternations pose a threat to public health.

***Keywords*---**SARS-CoV-2, Variants, Receptor binding domain, Spike glycoprotein, Genomics.

Introduction

In past, infectiuos disease with the potential to become pandemic have frequently developed and spread. Indeed, severe epidemics and pandemics have had an impact on humaity. SARS-CoV-2 is a menace to humanity on a global scale. This is the third epidemic of a very aggressive coronavirus capable of infecting people with serious pneumonia. All three coronavirus outbreaks occurred in the twenty-first century. The first pandemic of coronavirus occurred in the Chinese province of Guangdong in 2002 and continued until 2004. Severe acute respiratory syndrome coronavirus (SARS-CoV) was the catalyst for this event. The second epidemic was caused by the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) and was discovered in a male from Saudi Arabia in 2012 who had symptoms similar to pneumonia. The third most virulent coronaviral pandemic has been observed in China province Hubei, city Wuhan (Wong *et al.*, 2020). The first coronavirus found in 1937 was Infectious Bronchitis Virus (IBV) and was of poultry origin discovered by Beaudette & Hudson. The first Human coronavirus was discovered and revealed in the 1960s (Hasöksüz et al., 2020).

SARS-CoV-2 member to the order *Nidovirales*, family *Coronaviridae*, subfamily *Coronavirinae*, genus *betacoronavirus*. SARS-CoV-2 is a positive-sense, single-stranded RNA(+ssRNA), enveloped virus with about 30 kb of genome size. MERS-CoV-2 and SARS-CoV-2 share 50% and 80% of the genome, respectively (Kim *et al.*, 2020a; Kim *et al.*, 2020b). The subfamily

Coronavirinae have 4 genera *alphacoronavirus*, *betacoronavirus*, *deltacoronavirus* and *gammacoronavirus*. *Alpha*, *betacoronavirus* mainly infect mammals including humans, and *deltacoronavirus* and *gammacoronavirus* infect birds. HCoV-229E and HCoV-NL63 belongs to genus alphacoronavirus, while HCoV-HKU1, HCoV-OC43, MERS-CoV, and SARS-CoV belongs to genus *betacoronavirus*. All these are potential human pathogens and cause respiratory infections in humans (Corman *et al.*, 2020).

The virion of the coronavirus has four important structural proteins; Transmembrane (M) protein, Spike (S) protein, Envelop protein (E) and Nucleocapsid protein (N). These 4 proteins make the whole coronavirus. S protein makes up the spikes around the coronavirus. Because of S protein, the virus is known as coronavirus as it shows crown-like projections on its surface. S glycoprotein owned by the type 1 glycoproteins having varities of functional domains near amino (S1) and carboxy-terminal (S2). It is the S1 subunit that carryy the receptor-binding domain (RBD), whereas S2 is a transmembrane protein that is liable for mediating the fusion of virus with the membrane of the host cell. The overall structure of the S glycoprotein facilitates the virus binding to the host cell, cell fusion, and also neutralizes the antibodies. The S1 subunit has the higher activity of neutralizing. N protein has ability to form the nucleocapsid of virus and binds to the coronavirus genome. Nucleocapsid protein involves in genome-related processes of the virus like replication. Additionally, it is critical for the host's cellular response to the virus. M protein plays a major role as it uses the host membrane to make new virus particles. Moreover, it involves in the assembly of the virus. At the budding site, the interaction of M protein with the ribonucleoprotein (RNP) and S protein makes an M-M interaction network whose function is to remove the host membrane proteins from the viral envelop. E protein is the most peculiar structural protein. During the infection, although the cell produces large amounts of E protein, only a small amount of it is integrated into the envelope of virion (Hasöksüz *et al.*, 2020). Most of the E protein is located at the site of virus assembly, budding, intracellular trafficking, endoplasmic reticulum, and Endoplasmic reticulum Golgi intermediate components organelles. The main role of E protein is supposed to be its involvement in the virus pathogenesis, viral assembly, and hydrophobic transmembrane domain which is essential for the virion release (Schoeman *et al.*, 2019).

The genomic analysis of the novel coronavirus has indicated that, the most variable region in the spike protein is the key receptor binding domain (RBD). It is found that only 6 RBD amino acids are essential for ACE2 binding and determining the host range of SARS-CoV-like viruses. The analysis of SARS-CoV-2 structure has confirmed that RBD has a significant affinity for the ACE2 of human, cats, ferrets, and species with high receptor homology. Spike protein's strong affinity for human ACE2 (hACE2) receptors is an outcome of natural selection on human-like or human ACE2 receptors that would have permitted the optimal binding solution (Andersen *et al.*, 2020).

The mutation in the S region produced many variants. In SARS-CoV-2, a novel genetic mutation was confirmed in December 2020 in UK. It is referred

as variant of concern (VOC 202012/01) and belongs to the lineage B.1.1.7. N501Y mutation in the S region of the UK variant is responsible for the increased affinity towards the ACE2 receptors which facilitates viral attachment to the host cell and sucessive entrance into the host. This is the most dominant strain circulating in the world. A new variant of concern (B.1.351 lineage or 501Y.V2) with multiple mutations is reported in October 2020 in South Africa. This raised the likelihood of transmission and decreased the effectiveness of monoclonal antibodies, post-vaccination treatment, and convalescent sera in neutralizing the virus. The third variant of concern belongs to the B.1.1.28.1 lineage also known as P.1 lineage was found in December 2020 in Brazil. B.1.1.28.1 lineage may have minimized neutralization by monoclonal antibodies, post-vaccination therapy, and convalescent sera. Now, WHO have labeled the variants β, α, δ and γ for variants of concern and Eta, Iota, Kappa, Lambda and Mu for variants of interest (Cascella *et al.*, 2020). The present study was carried out to detect novel mutations in the RBD and SD1 region by amplifying the partial S gene by conventional PCR and the sequencing of this partial region.

Materials and Methods

Sample collection

The nasopharyngeal swabs in the sterile tube containing the virus transport medium (VTM) were collected from the suspected patients in the Medical Health Center, University of Agriculture, Faisalabad, Pakistan. During the collection and processing of the samples, all the precautionary measures were adopted to avoid the contamination and infection. All the work was done under Punjab health commission (PHC) certified BSL-II laboratory, UAF. This research was approved and monitored by the Institutional Biosafety committee, UAF.

Sample size

Samples from the suspected patients were collected in Virus Transport medium (VTM) in the Medical Health Center, University of Agriculture, Faisalabad. These suspected samples were transported to the lab.

Confirmation of SARS-CoV-2 through RT-qPCR

Before the verification of the SARS-CoV-2 through the RT-PCR, the RNA Extraction was carried out using the GeneJET Viral Purification Kit (Thermo Fisher Scientific, USA) according to provided guidelines. For the identification of SARS-CoV-2, SEEGENE Allplex SARS-CoV-2 kit was used. This kit detects 4 target genes by using RT-qPCR of SARS-CoV-2.

Conventional Polymerase Chain Reaction (PCR)

RNA cannot be amplified in conventional PCR. For this purpose, cDNA synthesis was carried out using ABM EasyScript Reverse transcriptase cDNA Synthesis Kit (Thermo Fisher Scientific, USA) according to guidelines. Spike gene region was amplified for sequencing purposes with polymerase chain reaction. The primers

that were used to amplify 719 bp of the S gene are shown in **Table 1.** Which included the SARS-CoV-2 Receptor binding domain or SD1 region (Kim *et al.*, 2020c). The PCR master mix was prepared containing the 25 µl master mix by ZOKEYO Internationals, 5 µl of the template cDNA, 1.5 µl from each corresponding primers and finally 17 µl of Nuclease free water was used to make the final volume upto 50 µl for each sample. Reaction conditions for the PCR were for 4 minutes at 94°C followed by 30 cycles for 30 seconds at 94°C, 30 seconds at 53.3°C, and for 45 seconds at 72°C. At final step extension was carried out at 72°C for 10 minutes. After the completion of the PCR run, 1% agarose gel was prepared containing the 1% ethidium bromide for the visualization. Finally, the amplified PCR product was run through Gel documentation system to separate and visualize the product on basis of their molecular weight to get the desired size of the S gene.

Table1: forward and reverse primers used for the detection of partial S gene region through PCR

Primer	Sequence
COV-2 Forward	5′- AACAATCTTGATTCTAAGGTTGGTGG -3′
COV-2 Reverse	5′- GTCTGAGTCTGATAACTAGCG -3′

Sequencing

Amplified product S gene was purified from the gel using Gel purification kit (ZOKEYO Internationals) according to the manufacturer's guidelines. The purified product containing the partial S gene was sent to the Zokeyo Internationals, China for DNA sequencing in order to ascertain the virus's nucleotide sequence. The Clustal W Method was utilized to align all the sequences using Lasergene's EditSeq and MegAlign tools.

Results

Ct-values of SARS-CoV-2

RT-qPCR was achieved using SEEGENE Allplex SARS-CoV-2 for all the samples collected on Bio-Rad CFX96. The ct values of all the twelve positive samples were analyzed using SEEGENE SARS-CoV-2 Viewer to interpret the results. The ct values of the positive samples were determined.

Conventional PCR

In accordance with the Material and Methods section, all samples that tested positive for SARS-CoV-2 were transformed into cDNA. Spike gene region was amplified for sequencing purposes with polymerase chain reaction kit by ThermoFishers scientific. PCR reaction was set up using the master mix recipe according to the manufacturer's protocol. During this whole procedure, all the delicate cDNA and other reagents were kept at -20°C for efficient results. Twelve COVID-19 patient's cDNA was employed as a template for amplification of the virus spike gene cDNA. PCR was utilized to produce a 719 bp band using a set of

forward and reverse primers. Out of these 12 samples only 3 samples have shown significance mutation.
Gel electrophoresis was used to separate the PCR products, and the presence of S gene was confirmed by examining the gel fragments. In 1X TBE buffer, a 1% agarose gel was produced. After the preparation of the gel, 6X loading dye was mixed with the amplified DNA and added to the wells of agarose gel along with known DNA markers to separate the bands. For 45 minutes, a current of 3 V/cm was delivered through the gel. Using a gel documentation system, the DNA bands were visible following Ethidium Bromide staining as illustrated in Figure 1.

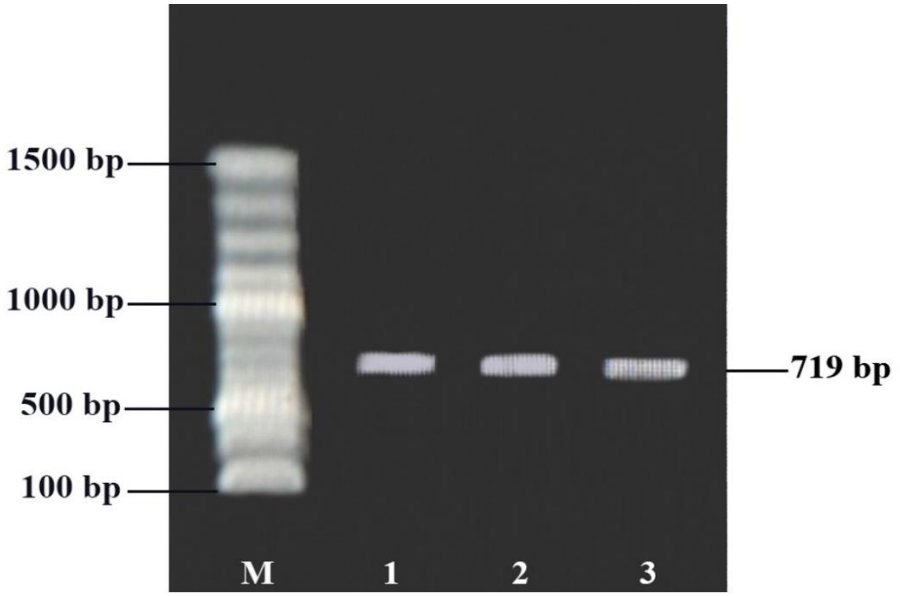

Figure 1: Represents S gene PCR Amplification. Lane M contains a 100 bp molecular marker. Lane 1-3 provides a 719 bp PCR results

Sequence analysis

Amplified PCR products were sent to ZOKEYO Internationals, China for sequencing. All the sequence were submitted to NCBI to gain the accession number. GenBank accession number are **MZ996568**, **OK001857**, and **OK001861**.

Construction of phylogenetic tree

Maximum Likelihood approach and Tamura-Nei model were used to determine the evolutionary history. It shows the tree with maximum log probability (-5101.98). The relevant taxa's percentage of clustered trees is displayed right next to the branches. The numbers along the branches represent the bootstrap values as a percentage of 1000 bootstrap resampling's. For the phylogeny, the presented samples were compared with the Human coronaviruses from genus Alpha and beta included the Wild-type SARS-CoV-2. There was an entire of 768 positions in the ultimate dataset containing the S gene of coronaviruses. Evolutionary

analysis were performed in MEGA X. SARS-CoV-2 samples are highlighted and labelled along with the accession numbers.

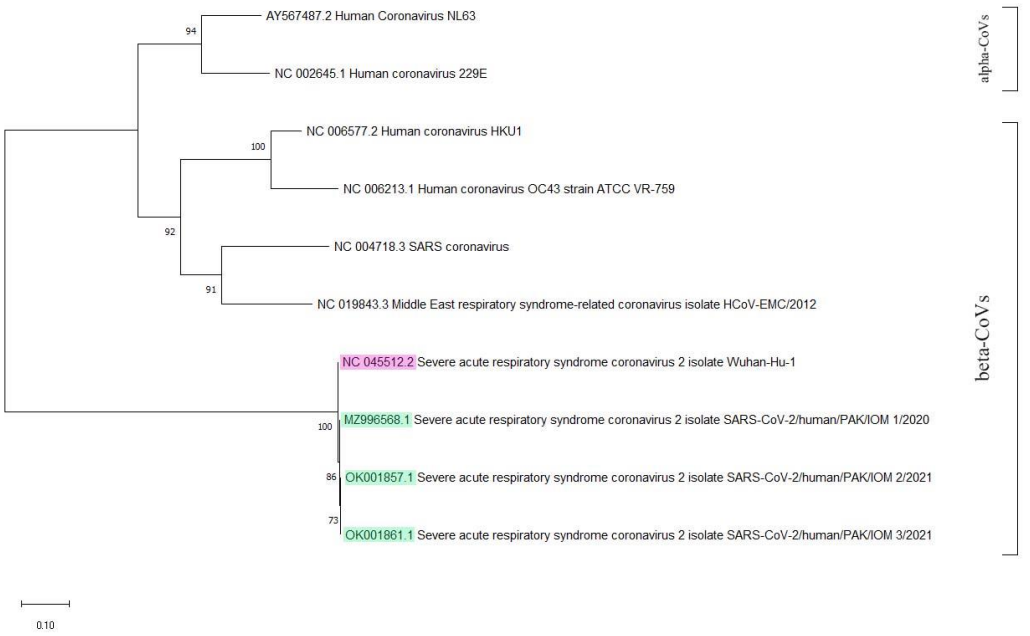

Figure 2: Inference of SARS-CoV-2 phylogenetic relationship with other coronaviruses. The SARS-CoV-2 wildtype is marked Pink while SARS-CoV-2 isolates IOM are marked with green. The remaining Alpha-CoVs and beta-CoVs are also mentioned in phylogenetic tree.

Mutation analysis

The mutation analysis was done using ClustalW method by comparing the obtained sequence of samples with wild type SARS-CoV-2. All samples have shown mutations at position 484 (E484K) where glutamic acid was substituted by lysine and 501 (N501Y) where asparagine was substituted by tyrosine. IOM 2 and IOM 3 have shown mutation at position 614 where aspartic acid substituted by glycine, but IOM 1 was not mutated at this site. In this study, the partial sequencing of S gene includes the RBD and some part of SD1 domains only. As indicated in Figure 3.

Figure 3: Partial 238 amino acid alignment indicating mutations highlighted with blue

E484K, N501Y and D614G mutations were seen in collected samples. **Figure 4** showing the entire genome and focusing on the spike gene, revealing important mutations that should be monitored in different parts of the world. The Spike gene's several sub-domains, which are shown by colored boxes, are described in the legend. Each line's position and color reflect the site of specific mutation, which is valued above that gene. The key mutational changes that were observed during this study are written with red letter. This graphical representation was inspired from (Srivastava *et al.*, 2021).

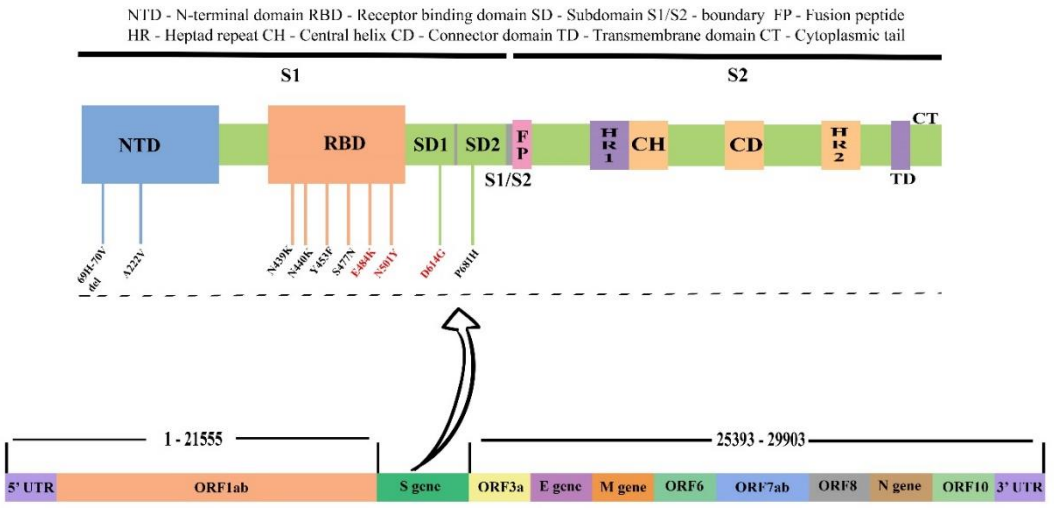

Figure 4: Graphical representation of whole genome with structural and non-structural regions, focusing on key mutations in S gene

Nucleic acid substitution

Sequencing data of the patient exhibited glutamic acid 484 codon “GAA” that was replaced by lysine 484 codon “AAA” and mutation was denoted as E484K. At position 501 codon “AAT” asparagine was replaced by tyrosine 501 codon “TAT” and mutation was denoted as N501Y. At position of codon 614 “GAT” aspartic acid which is replaced by codon 614 “GGT” of glycine and mutation was denoted as D614G.

Discussion

COVID-19 is brought on by SARS-CoV-2 belongs to betacoronavirus with a 30 kb single-strand RNA, positive-sense genome that encoding 29 key proteins. These include the structural proteins that virus uses to clump its RNA and the proteins that allow the virus to enter and spread inside the body by using host cellular machinery for viral reproduction. A lipid membrane produced by the host that surrounds the contained SARS-CoV-2 virus. The virus encodes numerous structural proteins, the most significant of which is the S protein, which forms a crown like composition around the viral envelope (Wu *et al.*, 2020). The Angiotensin-Converting Enzyme-2 (ACE2) receptor is located on numerous human cells, such as in nasal cavity, kidneys, intestines, lungs, heart, blood vessels and brain. A receptor binding domain is present in the Spike protein, which is produced by the S gene that has formed especially to attach with this receptor (Li *et al.*, 2020). Main entry of illness is through the oral and nasal route, where affected persons in close contact with healthy people spread virus particles that stick to the hosts' epithelial cells and allow virus to insert into the bodies. Numerous studies have discovered a link between an individual's level of ACE2 expression and the severity of the SARS-CoV-2 sickness, particularly in elderly

individuals, and individuals already having serious health issues (Wang *et al.*, 2020).

This study was carried out by collecting nasopharyngeal samples of suspected COVID-19 patients from the Medical Health Center, University of Agriculture Faisalabad. In current study, the screening of SARS-CoV-2 and to amplify the S gene region was carried out in order to detect any mutations. To prevent the virus particles from becoming denatured, the COVID-19 samples were collected in VTM tubes and transferred to the PHC accredited COVID-19 screening facility under controlled temperature conditions. The nucleic acid of all the samples were extracted by using the GeneJET Viral Purification Kit (Thermo Fisher Scientific, USA). The confirmation of SARS-CoV-2 in the unknown specimen were determined through RT-qPCR by using the SEEGENE Allplex assay to detect the genes specific to the 2019-nCoV. The threshold cycle values (ct values) in the samples were in the range of 23.27 to 38.19, suggesting a range of high- and low viral loads in the samples, respectively. The RNA from the positive samples was kept at -20°C until used for further testing and analysis (Kim *et al.*, 2020c).

(Korber *et al.*, 2020) recently revealed that the G614 strain is more contagious and have high upper respiratory tract virus burden; however, no correlation with disease severity was detected. When compared to the Wuhan reference strain, both groups had missense genetic mutations in spike gene; moreover, group S had a lower level of total diversity as compared to GH clade. This demonstrates that SARS-CoV-2 strains that are currently spreading in Pakistan have developed and also have different sequences than their ancestors. As a result, we advocate genome sequencing of indigenous strains found around the country in order to recognize strains with distinctive mutational patterns and to gain a better knowledge of virus evolution. The COVID-19 epidemic began at the end of 2019 and has wreaked havoc on economic, social, and medical systems in numerous countries. There isn't a specific COVID-19 medication or vaccine available right now, although various lines of inquiry are yielding encouraging results. It is essential to understand how SARS-CoV-2 enters host cells since disrupting this entry can prevent SARS-CoV2 multiplication and spread. It is very well established that the virus enters cells following digestion of the RBD from SARS-CoV-2 and S protein like viral envelope from host protease both these interact with receptor ACE2 (Wang *et al.*, 2020).

The D614G mutation altered the S protein's domain which is Receptor Binding Domain at 614th amino acid D (aspartic acid) with the amino acid G (glycine). Because glycine helps to create that has a hinge area that is more flexible in the Spike because it is a least nasty amino acid than aspartic acid, which allows more effective cutting for receptor. As a result, the virus had an advantage in terms of infection and transmission, allowing it to broad globally as studied (Zhang *et al.*, 2020). Other mutations such as E484K and N501Y were first detected in UK (VOC), and in lineages of South Africa and Brazil. P.1 lineage which is the descendent of B.1.1.28 have total 17-point mutation together with the E484K and N501Y. Several locations around Brazil, including Manaus, have revealed the existence of a unique sub lineage independent acquisition of the immune evasion-related spike E484K mutation in P.2. Recently, three cases of reinfection were found in Brazil, the P.1 lineage belonged to one, the P.2 lineage to the other two

(Nonaka *et al.*, 2021). These kinds of variants that have multiple point mutations can have high transmission rate or might escape the immune system which can enable them to reinfect the patients and can lower the vaccine efficacy worldwide.

Now it's been around 2 years since the pandemic of COVID-19. It is very necessary to have a surveillance on new, prevalent, and emergent strains. The change in the genetic makeup of the SARS-CoV-2 can make it highly contagious through aquired mutational changes leading to enhanced affinity and binding with the host cell. Some mutations can evade the immune system which will eventually have effect on vaccine efficacy and can lead to the reinfection of COVID-19. Many vaccines are being administered to the general public under the extended program on immunization by WHO. This will subject the virus under new selection pressure leading to the evolution (Srivastava *et al.*, 2021).

This preliminary study described multiple point mutations in the S gene that swapped the amino acid glutamic acid 484 with lysine (E484K), asparagine 501 with tyrosine (N501Y) and aspartic acid 614 with glycine (D614G) resulting in reduced sensitivity to antibody neutralization, increased affinity to receptor hACE2 for binding and enhanced transmission of the virus respectively. However, the D614G mutation's relevance is unknown because it was discovered within the SD2 gene, which has non-recognized specific role, as opposed to the RBD of the S gene. This research examined SARS-CoV-2 RBD of the S gene in COVID-19 victims from Pakistan. DNA sequencing study identified the E484K and N501Y mutations in three Pakistani COVID-19 cases and D614G mutation in two Pakistani COVID-19 sufferers, implying that the mutated cases of SARS-CoV-2 strain might have come out from the similar source.

Author Contributions

All authors have contributed to the final manuscript by division of responsibilities, i.e., literature search, compilation, write-up, final version development, English language editing, figures and table formation, formatting and reference styling, and proofreading. All authors have read and agreed to the published version of the manuscript.

Acknowledgement

The completion of this undertaking research could not have been possible without the participation and assistance of so many people whose names may not all be enumerated. Their contributions are sincerely appreciated and gratefully acknowledged. This work was supported by the Punjab Educational Endowment Fund (PEEF), Office of Research, Innovation & Commercialization (ORIC) and Sultan group. I would like to express my deep appreciation and indebtedness to InstaMed Medical Equipment Trading LLC. Other than that, no external fundings were received.

References

Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. 2020. The proximal origin of SARS-CoV-2. Nat. Med. 26: 450–452.

Cascella M, Rajnik M, Cuomo A, Dulebohn SC, Napoli RD. 2020 Features, Evaluation and Treatment Coronavirus (COVID-19), StatPearls Publishing.

Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, T, *et al.* 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 25: 2000045–2000053.

Hasöksüz M, Kiliç S, Saraç F. 2020. Coronaviruses and SARS-COV-2. Turkish J. Med. Sci. 50: 549–556.

Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. 2020a. The Architecture of SARS-CoV-2 Transcriptome. Cell. 181: 914-921.e10.

Kim JM, Chung YS, Jo HJ, Lee NJ, Kim MS, Woo SH, *et al.* 2020b. Identification of Coronavirus Isolated from a Patient in Korea with COVID-19 Osong Public Health and Research Perspectives. Public Heal. Res Perspect. 11 :3–7.

Kim S, Lee JH, Lee S, Shim S, Nguyen TT, Hwang J, Kim H, Choi YO, *et al.* 2020c. The progression of SARS coronavirus 2 (Sars-CoV2): Mutation in the receptor binding domain of spike gene. Immune Netw. 20: 1–11.

Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, *et al.* 2020 Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell. 182:812-827

Li MY, Li L, Zhang Y, Wang XS. 2020. Expression of the SARS-CoV-2 cell receptor gene ACE2 in a wide variety of human tissues. Infect. Dis. Poverty. 9: 1–7.

Nonaka CKV, Franco MM, Gräf T, de Lorenzo Barcia CA, Navesde Ávila Mendonça A, AF de Sousa K, *et al.* 2021. Genomic Evidence of SARS-CoV-2 Reinfection Involving E484K Spike Mutation, Brazil. Emerg. Infect. Dis. 2021. 27:1522–1524.

Schoeman D, Fielding BC. 2019. Coronavirus envelope protein: Current knowledge. Virol. J. 16.

Srivastava S, Banu S, Singh P, Sowpati DT, Mishra RK. 2021 SARS-CoV-2 genomics: An Indian perspective on sequencing viral variants. J. Biosci. 46.

Wong G, Bi YH, Wang QH, Chen XW, Zhang ZG, Yao YG. 2020 Zoonotic origins of human coronavirus 2019 (HCoV-19 / SARS-CoV-2): Why is this work important? Zool. Res. 41: 213–219.

Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, *et al.* 2020. A new coronavirus associated with human respiratory disease in China. Nat. 579: 265–269.

Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J. 2020. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. JAMA. 323: 1061-1069.

Zhang L, Jackson CB, Mou H, Ojha A, Peng H, Quinlan BD, *et al.* 2020. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. Nat. Commun. 11: 1–9.