Original Article

Comparative Analysis and Identification of Spike Mutations in Iranian COVID-19 Samples from the First Three Waves of Disease

Parisa Shoaei[1], Mohammad M. Ranjbar[2], Samaneh Tokhanbigli[3], Behrouz Ataei[3], Abbas Alibakhshi[4], Shaghayegh Haghjooy Javanmard[5], Shahrzad Ahangarzadeh[3]

[1]Nosocomial Infection Research Center, Isfahan University of Medical Sciences, Isfahan, Iran, [2]Razi Vaccine and Serum Research Institute, Agricultural Research, Education, and Extension Organization (AREEO), Karaj, Iran, [3]Infectious Diseases and Tropical Medicine Research Center, Isfahan University of Medical Sciences, Isfahan, Iran, [4]Molecular Medicine Research Center, Hamadan University of Medical Sciences, Hamadan, Iran, [5]Department of Physiology, Applied Physiology Research Center, Cardiovascular Research Institute, Isfahan University of Medical Sciences, Isfahan, Iran

Abstract

Background: The spike surface glycoprotein of SARS-CoV-2 is the essential protein in virus attachment to the target cell and cell entrance. As this protein contains immunodominant epitopes and is the main target for immune recognition, it is the critical target for vaccine and therapeutics development. In the current research, we analyzed the variability and mutations of the spike glycoprotein isolated from 72 COVID-19–positive patients from Iran's first three waves of disease.

Materials and Methods: The RNA was extracted from nasopharyngeal samples of confirmed COVID-19 cases and served as a template for cDNA synthesis and reverse transcriptase polymerase chain reaction. The reverse transcriptase polymerase chain reaction products of each sample were assembled and sequenced.

Results: After analysis of 72 sequences, we obtained 46 single nucleotide polymorphisms, including 23 that produce amino acid changes. Our analysis showed that the most frequent mutation was the D614G (in the samples of the second and third waves).

Conclusions: Our findings suggest that developing effective vaccines requires identifying the predominant variants of SARS-CoV-2 in each community.

Keywords: Mutation, SARS-CoV-2, spike glycoprotein

Address for correspondence: Dr. Shahrzad Ahangarzadeh, Infectious Diseases and Tropical Medicine Research Center, Isfahan University of Medical Sciences, Isfahan, Iran.
E-mail: shahrzadahangar@yahoo.com
Submitted: 24-05-2022; **Revised:** 16-11-2022; **Accepted:** 04-12-2022; **Published:** 28-Jun-2023

Introduction

Outbreaks and pandemics of infectious diseases have always profoundly affected different aspects of human life and living conditions. In the meantime, viruses have always played an important role in causing deadly epidemics.[1]

The COVID-19 pandemic, a major global health problem. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a positive-sense single-stranded RNA (+ssRNA) virus, has four structural proteins: spike (S), envelope (E), matrix (M), and nucleocapsid (N). The spike forms trimers on the virion's surface and composed of S1 and S2 subunits. The receptor-binding domain in S1 directly binds to the angiotensin-converting enzyme 2 (ACE2) receptor on the host cell surface. The S2 subunit is liable for virus-cell fusion.[2,3]

Access this article online

Quick Response Code:

Website: www.advbiores.net

DOI: 10.4103/abr.abr_171_22

How to cite this article: Shoaei P, Ranjbar MM, Tokhanbigli S, Ataei B, Alibakhshi A, Haghjooy Javanmard S, *et al*. Comparative analysis and identification of spike mutations in Iranian COVID -19 samples from the first three waves of disease. Adv Biomed Res 2023;12:153.

In addition to the critical role of spike protein in virus attachment and cell entrance, this protein contains immunodominant epitopes and is the main target for immune recognition. It is the primary purpose for vaccine and therapeutics development.[4,5] Single-nucleotide polymorphisms that cause amino acid changes in the surface proteins of viruses are worth monitoring because they can lead to phenotypic changes. Amino acid replacement in MERS-CoV[6] and SARS-CoV[7] made viruses resistant to neutralizing antibodies.

Gaining an insight into SARS-CoV-2 mutations, especially mutations in the key proteins, seems crucially important in evaluating virus behavior in pathogenesis, immune escape, and drug resistance.[8,9] Hence, we analyze the variability and mutations in this protein in isolated viruses from Iranian COVID-19 patients (in the first to third waves of the disease).

To achieve these, we used the extracted viral RNA from 72 samples of COVID-19–positive patients (from April 2020 to November 2020) for cDNA synthesis and amplification of fraction of spike gene (2,592 bp including S1 domain) by reverse transcriptase polymerase chain reaction (RT-PCR). After sequencing each fragment, the final sequences of each sample were assembled from three overlapping reads. We found 46 single nucleotide polymorphisms (SNPs), including 23 that produce amino acid changes. Evaluation of occurred mutations showed most of them have not yet been reported or described their roles.

Materials and Methods

Sample collection, RT-PCR, and sequencing

All the employed samples of COVID-positive patients (from Isfahan city, Iran) were obtained after informed consent in documents. Approval from the ethics committee was obtained in March 10, 2020. The approval ID is IR.MUI.MED.REC.1398.708. Afterward, extracted viral RNA from clinical samples confirmed by probe-based real-time PCR assay were selected for this study. For RT-PCR, the extracted RNAs were applied for cDNA synthesis as per the manufacturer's instruction for a cDNA synthesis kit (BioFACT, Korea). Three primer pairs[10] [Table 1] were used to amplify three consecutive fragments of the spike gene sequence. The amplification for the first fragment occurred with F1 and R1 primers at 94°C for 3.5 min, 94°C for 40 sec, 58°C for 45 sec, and 72°C for 2 min; 39 cycles and a final extension step of 72°C for 10 min. The second fragment was amplified with F2 and R2 primers (94°C for 3.5 min, 94°C for 40 sec, 59°C for 45 sec, and 72°C for 2 min; 39 cycles and a final extension step of 72°C for 10 min). The F3 and R3 primers were used to amplify the third fragment (94°C for 3.5 min, 94°C for 40 sec, 59.5°C for 45 sec, and 72°C for 2 min; 39 cycles and a final extension step of 72°C for 10 min).

Table 1: Sequences of designed specific primers for S1 and S2 subunits of SARS-CoV-2 spike

Primer	Sequence (5'-3')	PCR product size
F1	TATCTTGGCAAACCACGCGA	1,057 bp
R1	ACCAGCTGTCCAACCTGAAG	
F2	CCCTCAGGGTTTTTCGGCTT	1,190 bp
R2	CTGTGGATCACGGACAGCAT	
F3	CCAGCAACTGTTTGTGGACC	1,027 bp
R3	GTGGCAAAACAGTAAGGCCG	

RT-PCR products were sequenced using the Applied Biosystems SeqStudio Genetic Analyzer (ThermoFisher Scientific, Waltham, MA, USA). We used Finch TV and Bioedit software to confirm the sequencing quality (Geospiza Inc, Seattle, USA).

We assembled final sequences from three overlapping sequence reads and submitted all sequences to NCBI GenBank [Table 2].

Mutation analysis

To identify the SNP, we performed multiple sequence alignment of all 72 sequences using CLUSTALW implement of Bioedit 6.6.9 software with the S gene sequence of Wuhan as a reference sequence (Accession number: NC_045512). Nucleotide and amino acid sequences were analyzed using MEGA (molecular evolutionary genetics analysis) software (version X).[11]

Phylogenetic analysis

Multiple sequence alignment of S protein using CLUSTALW was performed for the phylogenic analysis of all 72 sequences and the required trimming was applied. A phylogenetic tree was developed by the maximum likelihood algorithm using MEGA X software.[11,12] The tree is drawn to scale and indicates the number of substitutions per site. The analysis involved 73 nucleotide sequences. Codon positions included were $1^{st}+2^{nd}+3^{rd}$+Noncoding. There were a total of 2,586 positions in the final dataset. Also, the evolutionary history was inferred by means of the maximum likelihood method as per the Kimura 2-parameter model.[12] The robustness of the tree topology was tested by bootstrapping (1,000 replicates).

Results

Identified single nucleotide and amino acid change mutations

We identified 46 mutations (the occurrence of one specific mutation in at least five samples was considered significant to report) and 23 of the detected mutations resulted in amino acid changes [Table 2]. Among these amino acid change mutations, T22I, G103R, D111N, Q115H, E224K, D228N, T240S, and A262T were in the NTD and L387I, Q474H, Q506H, and C525W were in receptor-binding domain. Four mutations, C538W, D614G, Q675R, and R685P, are in subdomains 1 and 2 [Figure 1a]. The most representative amino acid changes are substitution mutations such as E224K, D614G, D820N, V826G, and D830N.

The frequency of the G614 constantly increased from June to April 2020 and the cumulative frequency of the G614 was about 47.2% of all the analyzed sequences [Table 2]. The frequency of this mutation in the second-wave and third-wave samples was substantially increased compared to the first-wave samples [Figure 1b].

Table 2: Common nucleotide substitutions in 72 SARS-CoV-2 spike sequences (submitted to NCBI) compared to the SARS-CoV-2 NCBI reference genome (which were observed in more than 5 samples)

Nucleotide change	Amino acid change (if exists)	Synonymous/ Nonsynonymous	No. of SNPs ($n=72$)	NCBI Accession ID of the submitted sequence where the mutation is detected	%
G345C	Q115H	Nonsynonymous	9	MW321480.1 (45), MW322566.1 (50), MW114306.1 (63), MW114305.1 (66), MZ203839.1 (71), MZ203862.1 (82), MZ206020.1 (85), MZ206163.1 (91), MZ206172.1 (98)	12.5
G670A	E224K	Nonsynonymous	17	MW291652.1 (1), MW291693.1 (3), MW114660.1 (7), MW292556.1 (9), MW292557.1 (10), MW320665.1 (12), MW111529.1 (13), MW127506.1 (16), MW132908.1 (18), MW127548.1 (19), MW127502.1 (21), MW126756.1 (22), MW320670.1 (25), MW322030.1 (29), MW321481.1 (46), MW321483.1 (47), MW321496.1 (49)	23.6
G682A	D228N	Nonsynonymous	10	MW291652.1 (1), MW291693.1 (3), MW292557.1 (10), MW127528.1 (17), MW132908.1 (18), MW320670.1 (25), MW322030.1 (29), MW321481.1 (46), MW321496.1 (49), MZ203844.1 (76)	13.8
C882T		Synonymous	24	MW126756.1 (22), MW321496.1 (49), MZ203839.1 (71), MZ203840.1 (72), MZ203841.1 (73), MZ203842.1 (74), MZ203843.1 (75), MZ203857.1 (77), MZ203859.1 (79), MZ203860.1 (80), MZ203865.1 (81), MZ206017.1 (83), MZ206020.1 (85), MZ206021.1 (86), MZ206131.1 (87), MZ206132.1 (88), MZ206133.1 (89), MZ206163.1 (91), MZ206164.1 (92), MZ206166.1 (93), MZ206168.1 (95), MZ206170.1 (96), MZ206169.1 (97), MZ206174.1 (99)	33.33
A942G		Synonymous	6	MW321480.1 (45), MW321481.1 (46), MW322566.1 (50), MZ203842.1 (74), MZ203860.1 (80), MZ206133.1 (89)	8.33
A1290G		Synonymous	6	MZ203841.1 (73), MZ203842.1 (74), MZ203844.1 (76), MZ203860.1 (80), MZ206133.1 (89), MZ206166.1 (93),	8.33
T1446G		Synonymous	6	MW291652.1 (1), MW291651.1 (2), MW291693.1 (3), MW291959.1 (5), MW320665.1 (12), MW321483.1 (47)	8.33
C882T		Synonymous	24	MW126756.1 (22), MW321496.1 (49), MZ203839.1 (71), MZ203840.1 (72), MZ203841.1 (73), MZ203842.1 (74), MZ203843.1 (75), MZ203857.1 (77), MZ203859.1 (79), MZ203860.1 (80), MZ203865.1 (81), MZ206017.1 (83), MZ206020.1 (85), MZ206021.1 (86), MZ206131.1 (87), MZ206132.1 (88), MZ206133.1 (89), MZ206163.1 (91), MZ206164.1 (92), MZ206166.1 (93), MZ206168.1 (95), MZ206170.1 (96), MZ206169.1 (97), MZ206174.1 (99)	33.33
A942G		Synonymous	6	MW321480.1 (45), MW321481.1 (46), MW322566.1 (50), MZ203842.1 (74), MZ203860.1 (80), MZ206133.1 (89)	8.33
A1290G		Synonymous	6	MZ203841.1 (73), MZ203842.1 (74), MZ203844.1 (76), MZ203860.1 (80), MZ206133.1 (89), MZ206166.1 (93),	8.33
T1446G		Synonymous	6	MW291652.1 (1), MW291651.1 (2), MW291693.1 (3), MW291959.1 (5), MW320665.1 (12), MW321483.1 (47)	8.33
T1506G		Synonymous	9	MW291651.1 (2), MW291693.1 (3), MW291946.1 (4), MW127516.1 (8), MW320665.1 (12), MW127548.1 (19), MW321481.1 (46), MW321483.1 (47), MW322566.1 (50)	
G1662A		Synonymous	6	MZ203859.1 (79), MZ206017.1 (83), MZ206018.1 (84), MZ206020.1 (85), MZ206131.1 (87), MZ206163.1 (91),	8.33
A1841G	D614G	Nonsynonymous	34	MW291651.1 (2), MW320665.1 (12), MW116727.1 (15), MW126756.1 (22), MW321480.1 (45), MW321483.1 (47), MW321496.1 (49), MW132928.1 (65), MW114305.1 (66), MW114446.1 (69), MW114806.1 (70), MZ203839.1 (71), MZ203840.1 (72), MZ203841.1 (73), MZ203843.1 (75), MZ203857.1 (77), MZ203859.1 (79), MZ203865.1 (81), MZ203862.1 (82), MZ206017.1 (83), MZ206018.1 (84), MZ206020.1 (85), MZ206021.1 (86), MZ206131.1 (87), MZ206132.1 (88), MZ206133.1 (89), MZ206163.1 (91), MZ206164.1 (92), MZ206166.1 (93), MZ206167.1 (94), MZ206168.1 (95), MZ206170.1 (96), MZ206169.1 (97), MZ206172.1 (98)	47.2
G2054C	R685P	Nonsynonymous	6	MW321591.1 (57), MW321593.1 (58), MZ203842.1 (74), MZ203860.1 (80), MZ206133.1 (89), MZ206166.1 (93)	8.33

Contd...

Table 2: Contd...

Nucleotide change	Amino acid change (if exists)	Synonymous/ Nonsynonymous	No. of SNPs ($n=72$)	NCBI Accession ID of the submitted sequence where the mutation is detected	%
T2394G		Synonymous	9	MW291946.1 (4), MW291959.1 (5), MW292549.1 (6), MW292556.1 (9), MW292557.1 (10), MW110903.1 (11), MW320668.1 (24), MW320670.1 (25), MW320671.1 (26)	
T237C		Synonymous	6	MZ203859.1 (79), MZ206017.1 (83), MZ206018.1 (84), MZ206020.1 (85), MZ206131.1 (87), MZ206163.1 (91),	8.33
A2409C		Synonymous	6	MW291651.1 (2), MW291946.1 (4), MW320668.1 (24), MW320670.1 (25), MW320671.1 (26), MW322566.1 (50)	8.33
G2458A	D820N	Nonsynonymous	14	MW291652.1 (1), MW291651.1 (2), MW291946.1 (4), MW291959.1 (5), MW292549.1 (6), MW127528.1 (17), MW132908.1 (18), MW320666.1 (23), MW320668.1 (24), MW320670.1 (25), MW320673.1 (27), MW322030.1 (29), MW321481.1 (46), MW321496.1 (49)	19.44
T2477G	V826G	Nonsynonymous	11	MW291652.1 (1), MW291959.1 (5), MW127506.1 (16), MW127501.1 (20), MW126756.1 (22), MW320670.1 (25), MW320671.1 (26), MW320673.1 (27), MW321480.1 (45), MW321481.1 (46), MW321496.1 (49)	15.27
A2481C		Synonymous	12	MW291652.1 (1), MW291946.1 (4), MW291959.1 (5), MW127506.1 (16), MW127528.1 (17), MW127501.1 (20), MW320666.1 (23), MW320668.1 (24), MW320670.1 (25), MW320671.1 (26), MW320673.1 (27), MW321480.1 (45),	16.66
G2488A	D830N	Nonsynonymous	13	MW291652.1 (1), MW291946.1 (4), MW291959.1 (5), MW114660.1 (7), MW292557.1 (10), MW132908.1 (18), MW320668.1 (24), MW320670.1 (25), MW320671.1 (26), MW320673.1 (27), MW322030.1 (29), MW321496.1 (49), MW322566.1 (50)	18.05
T2514G		Synonymous	6	MW127506.1 (16), MW126756.1 (22), MW320670.1 (25), MW322030.1 (29), MW321483.1 (47), MW321496.1 (49)	8.33

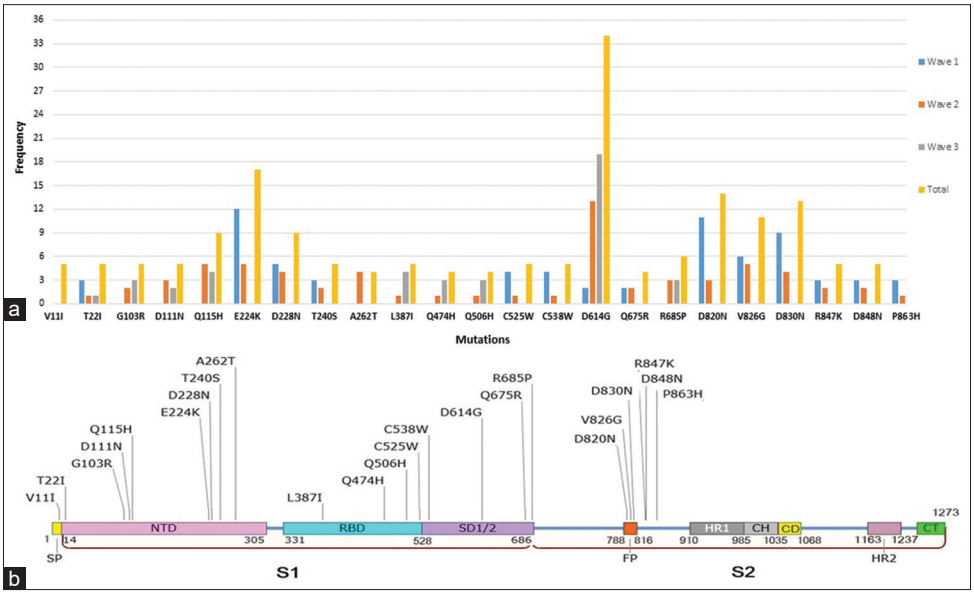

Figure 1: (a (Location of amino acid replacements in partial sequences of spike glycoprotein among the 72 sample sequences. (b) Frequency of amino acid replacements in sequences of spike glycoprotein

The spike D614G mutation could modify the spike structure, behavior, and infectivity of the virus. This mutation increases the affinity for ACE2[13] and may modify S1 domain binding activity.[14,15] In the samples of the first wave of COVID-19, E224K, D820N, and D830N owned the highest frequency, and in the second and third waves, D614G was the most frequent mutation in our samples [Table 3].

Phylogenetic analysis of spike gene

Figure 2 shows a phylogenetic tree illustrated with maximum likelihood (-8259.14). The percentage of trees in which the

Table 3: Nonsynonymous mutations were identified in the spike protein of 72 samples of SARS-CoV-2 in Isfahan, Iran, compared with hCoV-19/Wuhan/WIV04/2019

Amino acid Change	No. of nonsynonymous SNPs			
	Wave 1 *n*=24	Wave 2 *n*=24	Wave 3 *n*=24	Total *n*=72
V11I	0	5	0	5
T22I	3	1	1	5
G103R	0	2	3	5
D111N	0	3	2	5
Q115H	0	5	4	9
E224K	12	5	0	17
D228N	5	4	0	9
T240S	3	2	0	5
A262T	0	4	0	4
L387I	0	1	4	5
Q474H	0	1	4	4
Q506H	0	1	3	4
C525W	4	1	0	5
C538W	4	1	0	5
D614G	2	13	19	34
Q675R	2	2	0	4
R685P	0	3	3	6
D820N	11	3	0	14
V826G	6	5	0	11
D830N	9	4	0	13
R847K	3	2	0	5
D848N	3	2	0	5
P863H	3	1	0	4

associated taxa clustered together is displayed next to the branches.

The phylogenetic tree indicates the similarity of the sequence to each other. Generally, it presents three groups and eleven subgroups, and the first group, which includes the Wuhan strain, has the most subgroups. The sequences MZ203860, MZ203865, and MZ2033842 have more divergence and mutation than other sequences, respectively. Also, the sequence MZ206173 is the most similar and almost identical to the Wuhan sequence and has the slightest deviation from the ancestral node. Afterward, the sequences MW321691, MW321693, MW320677, MW111539, and MW127501 have the most similarity to the sequence of Wuhan, respectively [Figure 2].

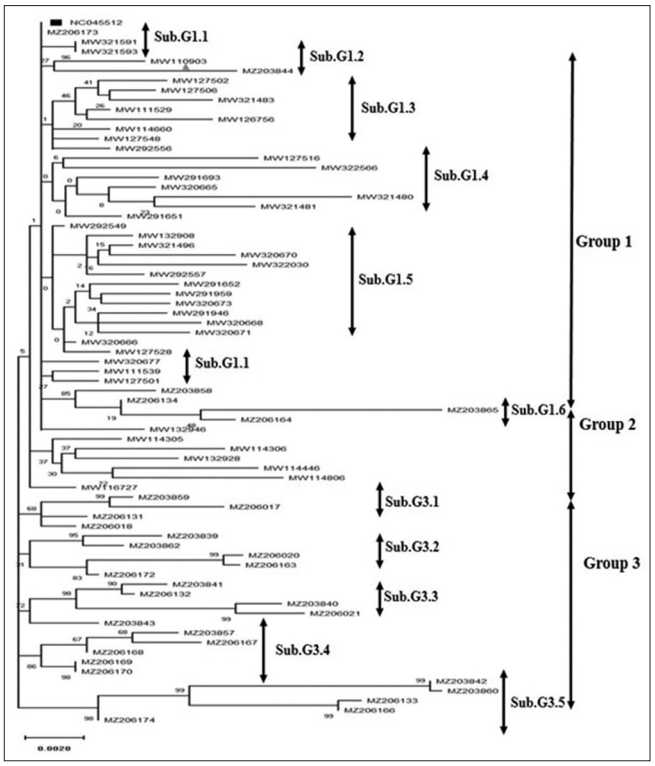

Figure 2: Phylogeny of SARS-CoV-2 spike glycoprotein by maximum likelihood method

Discussion

Genetic variations in SARS-CoV-2 could have important impacts on the severity and spread of disease. Reinfection, failure of the vaccination strategy, and failure of plasma therapy and monoclonal antibody therapy are also considered consequences of mutations in this virus. Some mutations also lead to the emergence of drug-resistant variants of the virus.[14-16] Accurately estimating the amount, type, and location of genetic diversity in the virus genome is essential for understanding the evolution of viruses and combating them, especially in the face of an epidemic and a virus with unknown characteristics. With the increase in cases of COVID-19, there are worries about the emergence of viral variants that are more severe and more toxic.

Spike protein is one of the prominent actors in the biology, pathogenicity, and immunogenicity of SARS-CoV-2. It is the basis of many anticoronavirus vaccines and one of the main targets in drug design.[17] Due to the high variability of the spike gene, it is necessary to consider mutations in prevention and control programs, preparation of polyvalent vaccine seed strains, and evaluation of drug resistance. Also, vaccination and antiviral treatments can lead to the creation of new mutations in circulating variants; therefore, genetic and clinical information on these variants is crucially important.[8,18]

Since Iran has no coherent and retrospective study on SARS-CoV-2 mutations, especially in the spike protein, this study was designed to evaluate the sequence, mutation pattern, and phylogenetic evaluation of the spike gene in the first three waves of the disease in Iran. Evaluation of occurred mutations showed most of them have not yet been reported or described their roles. Among the amino acid changes, the D614G mutation was the most frequent mutation in the samples of second and third waves. In a study conducted in Iran, Fattahi *et al*.[19] analyzed 50 whole-genome sequences of SARS-CoV-2 virus isolates from various geographical areas during the first wave of epidemics in Iran. The molecular and phylogenetic approaches showed that in the spike gene sequence, there were seven mutations in the spike protein. Meanwhile, two D614G and T22I mutations occurred at higher frequencies, as

observed in 17% and 24% of the samples, respectively. The rest of the mutations were seen in only one or two samples. As it has been shown, these two mutations also occurred in high frequencies in our samples, especially the D614G mutation in the second and third waves, in line with their findings, has shown the highest frequency.

These results could be helpful for the design and production of vaccines and therapeutics. It is worth mentioning that spike has been considered as the target in the design of recombinant and mRNA-based vaccines. The S1 subunit on spike protein contains the neutralizing antibody epitopes of the immune recognition sites; however, the mutations in this site have affected the efficiency of the candidate vaccines and demonstrate variation in different populations. Therefore, our findings suggest that developing effective vaccines requires identifying the predominant variants of SARS-CoV-2 in each community.

Acknowledgments

We gratefully acknowledge the Isfahan University of Medical Sciences, Isfahan, Iran, for providing financial and facility support (research grant no198231).

Declaration of patient consent

The authors certify that they have obtained all appropriate patient consent forms. In the form the patient(s) has/have given his/her/their consent for his/her/their images and other clinical information to be reported in the journal. The patients understand that their names and initials will not be published and due efforts will be made to conceal their identity, but anonymity cannot be guaranteed.

Financial support and sponsorship

The project was supported by Isfahan University of Medical Sciences [grant number: 198231].

Conflicts of interest

There are no conflicts of interest.

References

1. Nokhodian Z, Ranjbar MM, Nasri P, Kassaian N, Shoaei P, Vakili B, *et al*. Current status of COVID-19 pandemic; characteristics, diagnosis, prevention, and treatment. J Res Med Sci 2020;25:101.
2. Xue X, Shi J, Xu H, Qin Y, Yang Z, Feng S, *et al*. Dynamics of binding ability prediction between spike protein and human ACE2 reveals the adaptive strategy of SARS-CoV-2 in humans. Sci Rep 2021;11:1-12.
3. Long S, Olsen R, Christensen P, Bernard D, Davis J, Shukla M, *et al*. Molecular architecture of early dissemination and massive second wave of the SARS-CoV-2 virus in a major metropolitan area. mBio 2020;11:e02707-20.
4. Huang Y, Yang C, Xu X-F, Xu W, Liu S-W. Structural and functional properties of SARS-CoV-2 spike protein: Potential antivirus drug development for COVID-19. Acta Pharmacol Sin 2020;41:1141-9.
5. Kouhpayeh S, Shariati L, Boshtam M, Rahimmanesh I, Mirian M, Esmaeili Y, *et al*. The molecular basis of covid-19 pathogenesis, conventional and nanomedicine therapy. Int J Mol Sci 2021;22:5438.
6. Tang X-C, Agnihothram SS, Jiao Y, Stanhope J, Graham RL, Peterson EC, *et al*. Identification of human neutralizing antibodies against MERS-CoV and their role in virus adaptive evolution. Proc Natl Acad Sci 2014;111:E2018-26.
7. Mitsuki Y-Y, Ohnishi K, Takagi H, Oshima M, Yamamoto T, Mizukoshi F, *et al*. A single amino acid substitution in the S1 and S2 Spike protein domains determines the neutralization escape phenotype of SARS-CoV. Microbes Infect 2008;10:908-15.
8. Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, *et al*. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. Cell 2020;182:1284-94.e9.
9. Mohammadi E, Shafiee F, Shahzamani K, Ranjbar MM, Alibakhshi A, Ahangarzadeh S, *et al*. Novel and emerging mutations of SARS-CoV-2: Biomedical implications. Biomed Pharmacother 2021;139:111599.
10. Ren L-L, Wang Y-M, Wu Z-Q, Xiang Z-C, Guo L, Xu T, *et al*. Identification of a novel coronavirus causing severe pneumonia in human: A descriptive study. Chin Med J. 2020;133:1015-24.
11. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 2018;35:1547-9.
12. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980;16:111-20.
13. Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, *et al*. Spike mutation D614G alters SARS-CoV-2 fitness. Nature 2021;592:116-21.
14. Dimonte S, Babakir-Mina M, Taib Hama-Soor SA. Genetic variation and evolution of the 2019 novel coronavirus. Public Health Genomics 2021;24:54-66.
15. Jackson CB, Zhang L, Farzan M, Choe H. Functional importance of the D614G mutation in the SARS-CoV-2 spike protein. Biochem Biophys Res Commun 2021;538:108-15.
16. Koyama T, Platt D, Parida L. Variant analysis of SARS-CoV-2 genomes. Bull World Health Organ 2020;98:495-504.
17. Rees-Spear C, Muir L, Griffith SA, Heaney J, Aldon Y, Snitselaar JL, *et al*. The effect of spike mutations on SARS-CoV-2 neutralization. Cell Rep 2021;34:108890.
18. Xia X. Domains and functions of spike protein in sars-cov-2 in the context of vaccine design. Viruses 2021;13:109.
19. Fattahi Z, Mohseni M, Jalalvand K, Aghakhani Moghadam F, Ghaziasadi A, Keshavarzi F, *et al*. SARS-CoV-2 outbreak in Iran: The dynamics of the epidemic and evidence on two independent introductions. Transbound Emerg Dis 2022;69:1375-86.