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

Gene 813 (2022) 146113

Contents lists available at ScienceDirect

Gene

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

A comprehensive overview of identified mutations in SARS CoV-2 spike glycoprotein among Iranian patients

Solat Eslami [a,b], Mark C. Glassy [c], Soudeh Ghafouri-Fard [d,*]

[a] Department of Medical Biotechnology, School of Medicine, Alborz University of Medical Sciences, Karaj, Iran
[b] Dietary Supplements and Probiotic Research Center, Alborz University of Medical Sciences, Karaj, Iran
[c] Translational Neuro-Oncology Laboratory, San Diego (UCSD) Moores Cancer Center, University of California, CA, United States
[d] Department of Medical Genetics, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Edited by Xavier Carette

Keywords:
SARS CoV-2
Spike protein
Mutation
Variants of concern (VOC)
Variants of interest (VOI)

ABSTRACT

Since late 2019, when SARS-CoV-2 was reported at Wuhan, several sequence analyses have been performed and SARS-CoV-2 genome sequences have been submitted in various databases. Moreover, the impact of these variants on infectivity and response to neutralizing antibodies has been assessed. In the present study, we retrieved a total number of 176 complete and high-quality S glycoprotein sequences of Iranian SARS-COV-2 in public database of the GISAID and GenBank from April 2020 up to May 2021. Then, we identified the number of variables, singleton and parsimony informative sites at both gene and protein levels and discussed the possible functional consequences of important mutations on the infectivity and response to neutralizing antibodies. Phylogenetic tree was constructed to represent the relationship between Iranian SARS-COV2 and variants of concern (VOC), variants of interest (VOI) and reference sequence. We found that the four current VOCs – Alpha, Beta, Gamma and Delta – are circulated in different regions in Iran. The Delta variant is notably more transmissible than other variants, and is expected to become a dominant variant. However, some of the Delta variants in Iran carry an additional mutation, namely E1202Q in the HR2 subdomain that might confer an advantage to viral/cell membrane fusion process. We also observed some more common mutations such as an N-terminal domain (NTD) deletion at position I210 and P863H in fusion peptide-heptad repeat 1 span region in Iranian SARS-COV-2. The reported mutations in the current project have practical significance in prediction of disease spread as well as design of vaccines and drugs.

1. Introduction

Since December 2019, when the new coronavirus (severe acute respiratory syndrome coronavirus 2, SARS-CoV-2) was reported at Wuhan and quickly distributed all over the planet (Yao, 2020), several sequence analyses have been performed and SARS-CoV-2 genome sequences have been submitted in various databases (Cao, 2020). The Spike glycoprotein of this virus is a trimeric protein and each monomer comprises two functional subunits, S1 and S2 (Wang, 2020). The S1 subunit is involved in the receptor recognition and mediates attachment to the host cell receptor, while S2 subunit participates in the fusion of the virus with the cellular membranes (Lan, 2020). Mutations in SARS-CoV-2 spike-glycoprotein are of great interest as they mediate infection in human. Moreover, most of the approved vaccines are designed to induce immune responses against this protein (Schrörs, 2021). Therefore, continuous monitoring of SARS-CoV-2 spike-glycoprotein can provide valuable data, which can be helpful for early detection of variants that require adaptations in preventive and therapeutic strategies (Schrörs, 2021; Coutard, 2020).

Li et al. have assessed the influence of a number of variants and glycosylation site modifications on the infectivity of SARS-CoV-2 and their influence on response to neutralizing antibodies. They reported that S mutation D614G increases infectivity. While most of alterations in the receptor binding domain (RBD) reduce infectivity, A475V, L452R, V483A, and F490L variants induce resistance to some neutralizing antibodies (Li, 2020). More recently, Guruprasad has analyzed S protein

Abbreviations: GISAID, global initiative on sharing all influenza data; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; VOC, variant of concern; VOI, variant of interest.

* Corresponding author.
E-mail address: s.ghafourifard@sbmu.ac.ir (S. Ghafouri-Fard).

https://doi.org/10.1016/j.gene.2021.146113
Received 22 July 2021; Received in revised form 31 October 2021; Accepted 6 December 2021
Available online 9 December 2021

sequences from almost all over the World and compared these sequences with the reference sequence from Wuhan-Hu-1. He has reported that about 80% of analyzed protein sequences contain at least one mutation. The observed mutations have been shown to occur in 400 distinctive mutation sites. Notably, 44 mutations have been reported in the RBD which participate in the interplays with the ACE2 receptor causing COVID-19 disease (Guruprasad, 2021). It is known that variation in the epitopes of SARS-CoV-2 RBD over time might introduce novel mutations which might challenge the development of broadly protecting antibodies and vaccines against SARS-CoV-2 (Sun, 2020). These mutational variants might also lead to escape from the pre-existing cross-reactive CD4+ T cell responses or long-term protection from re-infection through T cell memory (Braun, 2020). Hence, there is a necessity of constant monitoring of the rapidly changing mutation rates in the spike protein in SARS-CoV-2, which could have significant impact on virus infection, transmissibility and pathogenicity in the current pandemic (Schrörs, 2021).

In the present study, we have retrieved sequences of S protein of SARS-CoV-2 viruses detected in Iran to identify notable mutations of S protein compared to the Wuhan isolate and representative sequences of variants of concern (VOC) and variants of interest (VOI). We also discuss the observed S protein mutations in terms of their distribution in different regions of Iran as well as sites and types of mutations. The reported mutations in the current project have practical significance in prediction of disease spread as well as development of vaccines and drugs.

2. Materials and methods

2.1. Assortment S glycoprotein gene data and processing

Genetic variations of the S glycoprotein were identified in Iranian SARS-CoV-2 strains through the following steps. All SARS-CoV-2 sequences from Iran were obtained from the global initiative on sharing all influenza data (GISAID) and GenBank (National Center for Biotechnology Information, USA) submitted from April 2020 up to 24 May 2021. We also downloaded the representative sequences of S glycoproteins of VOC and VOI from GISAID for comparison with the S glycoproteins of Iranian SARS-COV-2. These sequences were subsequently aligned to the reference SARS-CoV-2 sequence (NC_045512). This step was accomplished using the MUSCLE program (Edgar, 2004) implemented in Mega X (Kumar, 2018) for amino acid or nucleotide sequences. The complete genome sequence of SARS-CoV-2 Wuhan-Hu-1 strain (NC_045512.2) was considered as the reference. The sequences that did not cover the entire *S* region were excluded. Furthermore, we excluded sequences that comprised the nucleotide ambiguity (N) and low-quality sequences. We used only one sequence for samples that have been repeated in both GenBank and GISAID. The final dataset included 176 complete and high-quality S glycoprotein sequences that were gathered during April 2020 until May 2021.

2.2. Identification of the S gene mutations

The complete SARS-CoV-2 *S* gene sequences were analyzed using the Molecular Evolutionary Genetics Analysis software (MEGA X) (Kumar, 2018). The aligned amino acid sequences were visualized. Mutations were recognized through comparing the sequences with the reference SARS-CoV-2 sequence (NC_045512), which was regarded as the wild-type. Amino acids that were translated from codons comprising ambiguous bases (e.g. R, Y) were removed from mutation analyses.

The number of variables, singleton, and parsimony informative sites were tabulated at both gene and protein levels. After filtering the ambiguous data, the amino acid substitutions in S glycoprotein of all Iranian SARS-CoV-2 samples were determined. Residues that exhibited a mutation in at least 5 strains were considered as frequently mutated residues.

Table 1
The results of nucleotide and protein alignment of 176 *S* gene of SARS-COV-2 from Iranian patients.

Name	Total number of samples	Number of nucleotides	Variable sites	Variable sites per 100 bp	Synonymous mutations	Non synonymous mutations	Singleton informative sites per 100 bp	Parsimony informative sites per 100 bp	Number of amino acids	Variable site	Variable site per 100 amino acids	Singleton informative sites per 100 amino acids	Parsimony informative sites per 100 amino acids
Spike protein	176	3822	450	11.77	192	258	9	2.77	1274	258	20.25	16.09	6.27

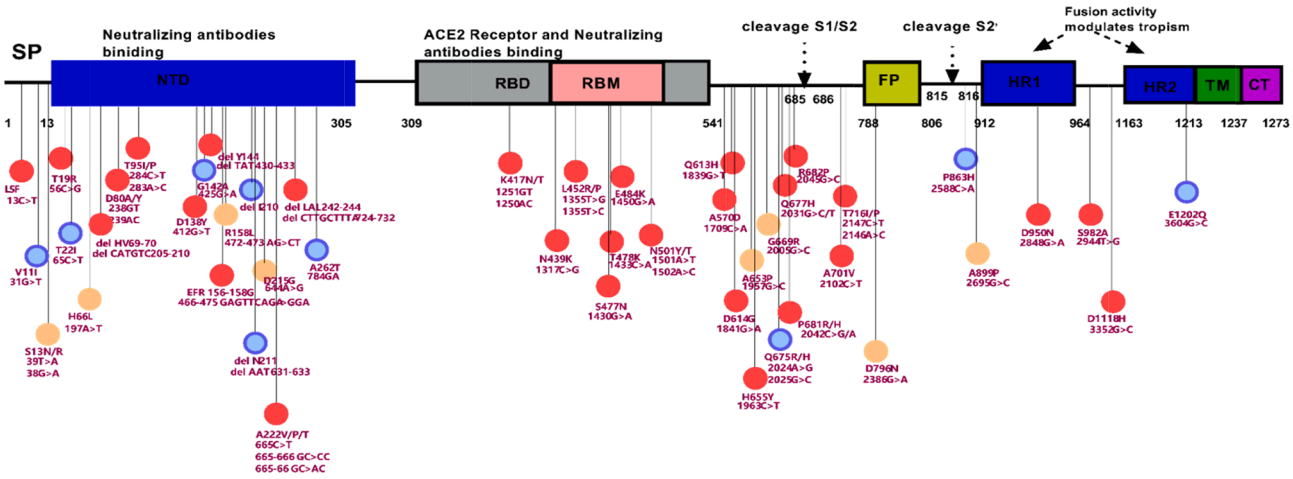

Fig. 1. Mutational profile of Iranian SARS-CoV-2 spike proteins. Commonly mutated residues, i.e. those observed in at least 5 SARS-CoV-2 isolates, specific mutations as indicator of variants of concern (VOC) and variants of interest (VOI) are shown on the respective regions of the SARS-CoV-2 spike gene. Mutations as indicator of VOC and VOI are depicted with reddish and orange (substitution mutations with other amino acids) circles, and the frequent uncommon mutations are depicted with blue circles. Domain organization of spike: Signal peptide: 1–13aa; N-terminal domain (NTD): 13–305aa; Receptor binding domain (RBD): 319–541aa; Receptor binding motif (ACE2):437–508aa; Furin cleavage sequence: 680–685aa; Fusion peptide (FP): 788–806aa; Heptad repeat region 1 (HR1): 912–964aa; Heptad repeat region 2 (HR2): 1163–1213aa; Transmembrane domain (TMD): 1214–1237aa; Cytoplasmic domain: 1238–1273aa. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.3. Phylogenetic analysis

The acquired S gene sequences from Iranian SARS-COV-2 strains were aligned with representative *S* gene of SARS-CoV-2 sequences of the eleven recognized GISAID variants (4 VOCs and 7 VOIs) publicly available in the GISAID database and the Wuhan reference strain, (NC_045512) using MUSCLE multiple sequence alignment algorithms executed in MEGAX (Edgar, 2004; Kumar, 2018). For the purpose of simplicity, and to avoid clutter, of the 176 S gene sequences, only distinct S gene sequences (131 sequences) along with 11 sequences of VOC and VOI and one reference sequence were included for phylogenetic tree analysis. The evolutionary history was deduced by using the Maximum Likelihood method and Tamura-Nei model (Tamura and Nei, 1993). The tree topologies were supported by 1000 bootstrap replicates. The studied S gene of Iranian SARS-COV-2 was dispersed into many subbranches and some sequences were clustered in closer with the respective representative S gene of VOCs and reference sequence.

2.4. The clade classification of Iranian SARS-COV-2 strains

The viral clade distribution of Iranian SARS-COV-2 during the pandemic time were determined according to the information about Iranian SARS-CoV-2 complete genome sequences submitted in GISAID. A total number of 267 whole genome sequences of Iranian SARS-COV-2 has been deposited in public database of GISAID from April 2020 up to May 2021. However, some of these 267 SARS-COV-2 sequences had nucleotide ambiguity (N) and low-quality sequences in *S* gene sequences and some had identical *S* gene sequences. Thus, these sequences were excluded from mutation and phylogenetic tree analyses; however, the GISAID data of all of sequences were used for clades distribution analysis of SARS-CoV-2 in Iran. Dynamics and diversity of viral clades were identified during the pandemic time in Iran.

3. Results

We retrieved a total number of 176 complete and high-quality S glycoprotein sequences of Iranian SARS-COV-2 submitted in public database of the GISAID and GenBank from April 2020 up to May 2021. Supplementary Table S1 shows the mutation sites and mutation types detected in human SARS-CoV-2 spike proteins from different regions of Iran. Supplementary Table S2 shows the list of accession numbers and date of collection of SARS-COV-2 strains from Iran (176 sequences) and representative of VOC and VOI from GISAID (11 sequences) and Wuhan-Hu-1(NC_045512.2) as reference. The complete length of the S glycoprotein sequences of human SARS-CoV-2 is 3822 nt, of which 450 variable sites were identified in 176 S glycoprotein sequences. Of these variable sites, 106 were parsimony informative, while 344 were singleton mutations. The results of nucleotide and protein alignment of SARS-CoV-2 *S* gene sampled from Iranian patients are shown in Table 1.

The identified mutations were distributed in different domains of spike protein with higher numbers of mutations being detected in NTD possibly resulting in escape from neutralizing antibodies. Fig. 1 depicts the mutational profile of Iranian SARS-CoV-2 spike proteins.

Table 2 shows distribution of mutations (commonly mutated residues and specific mutations as indicator of VOCs and VOIs) among different domains of S protein. This table also shows the number of occurrences of wild type and mutant residues among the sequences of Iranian SARS-CoV-2 spike genes.

Next, we retrieved the most important mutations in signal peptide and N-terminal domain (NTD: 13–305aa), the number of events in Iranian spike protein sequences and assessed their possible functional implications based on the available reports in the literature (Table 3).

Similarly, mutations observed in RBD were assessed in terms of their implications in infectivity and clinical responses (Table 4).

Moreover, we retrieved list of important mutations in RBD-Fusion peptide span region (Table 5), and those within Fusion peptide domain, Fusion Peptide- Heptad repeat 1span region, Heptad repeat 1, Heptad repeat 1- Heptad repeat 2 span region and Heptad repeat 2 (Table 6), as well as their clinical significance.

Subsequently, we depicted the phylogenetic tree of SARS-CoV2 S gene sequences in Iran based on spike gene (Fig. 2). In the phylogenetic tree, the studied *S* gene is dispersed into many subbranches and some sequences are clustered close with the respective representative S gene of VOCs and reference sequence. we attempted to determine the amino acid mutations discriminating the sequences into different clusters based on their similarities. Amino acid mutation or combination of them had a good discrimination ability, and clearly indicated the phylogenetic relationship between SARS-COV-2 strains and their association with VOCs, VOIs and the reference strain. The phylogenetic analysis clearly clustered the possible VOC of Delta, Beta, Gamma and Alpha of Iranian

Table 2
Distribution of important mutations among different domains of S protein (Signal peptide, N-terminal domain (NTD), Receptor binding domain (RBD), Receptor binding domain-Furin peptide linker; Fusion peptide (FP), Heptad repeat region 1 (HR1), HR1-HR2 linker, Heptad repeat region 2 (HR2).

protein	Spike protein																			
Protein domain	Signal peptide			N terminal domain (NTD)																
AA position	5	11	13	19	22	66	69-70	80	95	138	142	144	156-158	158	210	211	215	222	242-244	262
WT (#)	L (175)	V (171)	S (173)	T (167)	T (171)	H (175)	HV (168)	D)171(	T (169)	D (171)	G (166)	Y (169)	EFR (167)	R (174)	I (165)	N (173)	D (175)	A (171)	LAL (175)	A (160)
Mutant (#)	F (1)	I (5)	N (2), R (1)	R (9)	I (5)	L (1)	- - (8)	Y (4), A (1)	I (6), P (1)	Y (5)	D (8), A/S (2)	- (7)	G (9)	L (2)	-(11)	- (1), Y (2)	G (1)	V (3), P/T (2)	- - - (1)	T (15), P (1)

protein																									
Protein domain	RBD						RBD-FP linker												FP	FP-HR1 Linker		HR1	HR1-HR2 Linker		HR2
AA position	417	452	477	478	484	501	570	613	614	653	655	669	675	677	681	682	701	716	796	863	899	950	982	1118	1202
WT (#)	K (174)	L (166)	S (172)	T (167)	E (173)	N (167)	A (172)	Q (175)	D (98)	A (174)	H (173)	G (175)	Q (171)	Q (172)	P (162)	R (175)	A (175)	T (171)	D (175)	P (166)	A (166)	D (164)	S (172)	D (172)	E (172)
Mutant (#)	N/T (2)	R (9), P (1)	N (4)	K (9)	K (2), Q (1)	Y (9)	D (4)	H (1)	G (78)	P (2)	Y (3)	R (1)	R (4), H (1)	H (4)	R (9), H (5)	P (1)	V (1)	I (4), P (1)	N (1)	H (10)	P (1)	N (12)	A (4)	H (4)	Q (4)

(#) number of occurrences of wild type and mutant residues
(#) number of occurrences of wild type and mutant residues.

SARS-COV-2 in closer with the respective representative S gene of VOCs used in this study.

Nine SARS-CoV-2 variants in Iran were classified as Delta variant. These variants were reported from Yazd, Maku and Bushehr from April 2021 up to May 2021 (Table 7). Four of these nine sequences sampled from Yazd, in addition to indicator mutations of a Delta variant, also have a common E1202Q mutation in the HR2 subdomain, in which glutamine replaces glutamic acid at position 1202.

A total of 11 variants in Iran carry I210del mutation. These variants were reported from Shiraz, Tehran, Gilan and Urmia (Table 8).

We also found a total of seven sequences in Iran carrying all or some of specific mutations as indicator of Alpha variant. These sequences were reported from Shiraz, Tehran and Kerman (Table 9).

There were also three sequences carrying D138Y + S477N + D614G triple mutations sampled from shiraz and Tehran. the Main characteristics of these three SARS-CoV-2 variants are shown in Table 10.

There were two sequences carrying all of specific mutations as indicator of Beta variant. These sequences were reported from Hormozgan (Table 11). The Main characteristics of these SARS-CoV-2 variants are shown in Table 11

We found a SARS-COV-2 spike protein sequence sampled from shiraz carrying five specific mutations as indicator of Gamma variant. Table 12

Table 3
List of important mutations observed in signal peptide and N-terminal domain (NTD: 13–305aa), events and possible functional implications among spike protein sequences submitted in NCBI and GISAID from Iranian SARS-COV-2 strains.

Sr. No	Mutations	Events	Possible functional implications
1	L5F	1	Increases the epitope binding affinity for CD8 cell (Guo and Guo, 2020)
2	T19R	9	Might disrupt the "super site" on the NTD mediating neutralization (Liu, 2021)
3	H49Y	1	Increases the stability of spike glycoprotein (Laha, 2020)
4	L54F	1	When coupled with D614G, increases the stability of spike glycoprotein (Laha, 2020)
5	del HV69-70	8	Enhances viral infectivity and could also contribute to antibody evasion (Kemp et al., 2021)
6	D80A	1	Contributes to neutralization escape (Wibmer, 2021)
7	T95I	6	Might adversely affect the neutralization by antibody (West et al., 2021)
8	D138Y	5	Disrupts the epitope for mAb (Dejnirattisai, 2021)
9	G142D	8	Confers resistance to mAb (Suryadevara, 2021)
10	Del Y144	7	Confers resistance to mAb (Suryadevara, 2021)
11	EFR156-158G	9	Immune evasion through antibody escape (Chaudhari et al., 2021)
12	Del I210	11	ND
13	D215G	1	Contributes to neutralization escape (Wibmer, 2021)
14	del LAL 242–244	1	Confers resistance to mAb (Wang, 2021)
15	R246I	1	Confers resistance to mAb (Wang, 2021)

Table 4
List of important mutations observed in Receptor binding domain (RBD): 319–541aa, events and possible functional implications among spike protein sequences submitted in NCBI and GISAID from Iranian SARS-COV-2 strains. Mutations in Receptor binding motif (RBM): 437–508aa has been showed with stars.

Sr. No	Mutations	Events	Possible functional implications
1	R408K	1	Attenuates monoclonal and serum antibody neutralization (Liu et al., 2020)
2	K417N/T	2	Increases binding of the spike RBD to the ACE2 and higher infectivity (Khan, 2021)
3	N439K*	3	Enhances binding affinity for human ACE2 and confers resistance to numerous neutralizing monoclonal antibodies and permits immune escape from some polyclonal sera (Flemming, 2021)
4	L452R*	9	Evades cellular immunity and increases infectivity (Motozono, 2021; Li, 2020)
5	S459Y*	1	Might increase binding of the spike RBD to the ACE2 (Wang, 2021)
6	N460I*	1	Could escape some monoclonal antibody (Starr, 2021)
7	S477N*	4	Fortifies the binding of the SARS-COV-2 spike with the hACE2 receptor (Singh, 2021)
8	T478K*	9	Increases the binding of the spike RBD to the ACE2 and escape immune recognition (Di Giacomo, 2021; Wang, 2021)
9	E484K*	2	Reduces the binding of serum polyclonal neutralizing antibodies (Jangra, 2021; Harvey, 2021)
10	G485S*	1	Decreases the ability of the S protein to bind ACE2 (Zhang et al., 2021)
11	N501Y*	9	Enhanced binding of SARS-CoV-2 spike protein to the human ACE2 receptor (Luan et al., 2021)
12	H519Y	1	Replacement of histidine with proline in this position increases resistance to human covalescent sera neutralization (Li, 2020)

Table 5
List of important mutations observed in RBD-Fusion peptide span region, events and possible functional implications among spike protein sequences from Iranian SARS-COV-2 strains. Mutations in Furin cleavage site: 680–685aa has been shown with stars.

Sr. No	Mutations	Events	Possible functional implications
1	A570D	4	Might contribute to immune evasion of variants (Lazarevic, 2021)
2	D614G	78	Increases viral replication in the upper respiratory tract and enhances the vulnerability of the virus to neutralization by antibodies (Plante, 2021; Zhang, 2020)
3	H655Y	3	Confers escape from human monoclonal antibodies (Braun, 2021)
4	Q675H	1	Might confers immune escape (Grabowski, 2021) and perhaps affects the cleavage of RRAR, a critical step for virus entry (Zhou et al., 2020)
5	Q675R	4	Possibly influences the cleavage of RRAR, an important step for virus entry (Zhou et al., 2020)
6	Q677H	4	Might influence cell entry due to its closeness to the polybasic cleavage site (Zhou et al., 2020; Tomkins-Tinch et al.)
7	P681R*	9	Confers the neutralizing antibody resistance, enables the furin-mediated spike cleavage and increases cell–cell fusion (Saito et al., 2021)
8	P681H*	5	May increase spike cleavage by furin-like proteases, this does not significantly impact viral entry or cell–cell spread (Lubinski et al., 2021)
9	A701V	1	Might decrease the neutralization capacity of antibodies provoked by infection with preceding variants or vaccination (Lazarevic, 2021)
10	T716I	4	Might be associated with increased transmissibility and potential immune evasion (Lazarevic, 2021)

Table 6
List of important mutations observed in Fusion peptide domain, Fusion Peptide-Heptad repeat 1span region, Heptad repeat 1, Heptad repeat 1- Heptad repeat 2 span region and Heptad repeat 2, events and possible functional implications among spike protein sequences submitted in NCBI and GISAID from Iranian SARS-COV-2 strains.

Sr. No	Mutations	Events	Possible functional implications
1	D796N (FP)	1	The spike substitution mutant D796H seems to contribute to the decreased susceptibility to neutralizing antibodies (Kemp, 2021), however, the possible functional implications of the substitution mutant D796N is yet unknown.
11	P863H (FP-HR1 span)	10	ND
12	D936Y (HR1)	1	Strongly destabilizes the post-fusion configuration, while having a borderline effect on the stability of the pre-fusion one (Cavallo and Oliva, 2020)
13	D950N (HR1)	12	Viral Oligomerization Interfaces (Mishra et al., 2021)
14	S982A (HR1-HR2 span)	4	This mutation expands in lineage (B.1.1.7) and might be associated with relatively high infectivity (Kemp et al., 2021).
15	D1118H (HR1-HR2 span)	4	This mutation expands in lineage (B.1.1.7) and might be associated with relatively high infectivity (Kemp et al., 2021).

shows the main characteristics of the variant having specific mutations of a Gamma variant in Iran.

Finally, main characteristics of the fifteen SARS-CoV-2 variants carrying A262T mutation in Iran are shown in Table 13.

We also depicted the clades distribution of Iranian SARS-CoV-2 samples over time. Dynamics and diversity of viral clades are shown during the pandemic time in Iran (Fig. 3). There was a total number of 267 whole genome sequencing of Iranian SARS-COV-2 in public database of the GISAID from April 2020 up to May 2021.

There was an initial period between February and June 2020 when the O clade (GISAID nomenclatures) was more prevalent in total than other clades (Fig. 3b). The clade O then slowly declined and almost disappeared in early November 2020, while the GISAID GH clade peaked at ca 65% in August and September 2020 and has slowly declined since December 2020. The GISAID GR clade peaked at ca. 40% in October 2020 up to February 2021 and then rapidly declined. The GISAID GRY clade slowly increased from January 2021 and almost peaked at ca. 70% in March 2021. The clade G has been from May 2020 up to February 2021 under 20% of samples, however it peaked at ca. 50% in April 2021.

The GISAID clades are distinguished according to nucleotide variants that produce amino acid changes in different genes. According to last information about Iranian SARS-CoV-2 complete genome sequences submitted in GISAID, currently more than 95% of the Iranian strains fit to the two main branches of the G clade, i.e., GR and GH clades and the newly introduced GRY clade. The shared characteristic feature of these three clades is S_D614G mutation, which may increase the infectivity of SARS-CoV-2 (Brufsky, 2020).

4. Discussion

spike (S) protein is closely implicated in the instigation of SARS-CoV-2 infection. This protein is also the main target of neutralizing antibodies; thus, it is evolving with a high speed. Besides, alterations in S protein explain at least a number of phenotypic variations expressed by VOC (Liu, 2021). All VOCs (B.1.1.7, B.1.351, and P.1) and three B.1.617 sublineages have exhibited mutations inside and near the ACE2-interacting surface of S protein (Liu, 2021).

We showed that the SARS-CoV-2 S protein contains tens of mutations in Iranian patients when compared to the initial SARS-CoV-2 sequence from Wuhan. These mutations are distributed in different domains of spike protein with higher numbers of mutations being detected in NTD, possibly resulting in escape from neutralizing antibodies. L452R, T478K, N501Y and S477N have been the most frequently detected mutations in RDB. Notably, N501Y and S477N has also been reported as frequent mutations in this domain in a World-wide assay (Guruprasad, 2021). Moreover, N501Y mutation is among mutations which are located at the boundary between the spike protein and ACE-2 receptor and can influence efficiency of vaccines and drugs targeted the interface of protein–protein interactions. In NTD, the most commonly detected mutations have been Del I210, EFR156-158G and T19R with the latter possibly disrupting the “super site” on the NTD (Liu, 2021). Moreover, EFR156-158G has a role in immune evasion through antibody escape (Chaudhari et al., 2021). The D614G mutation as the only frequent mutation in the spike protein detected up to now among all the continents (Guruprasad, 2021), has been the most commonly detected mutation in RBD-Fusion peptide span region in Iranian samples.

Four of nine sequences sampled from Yazd, in addition to indicator

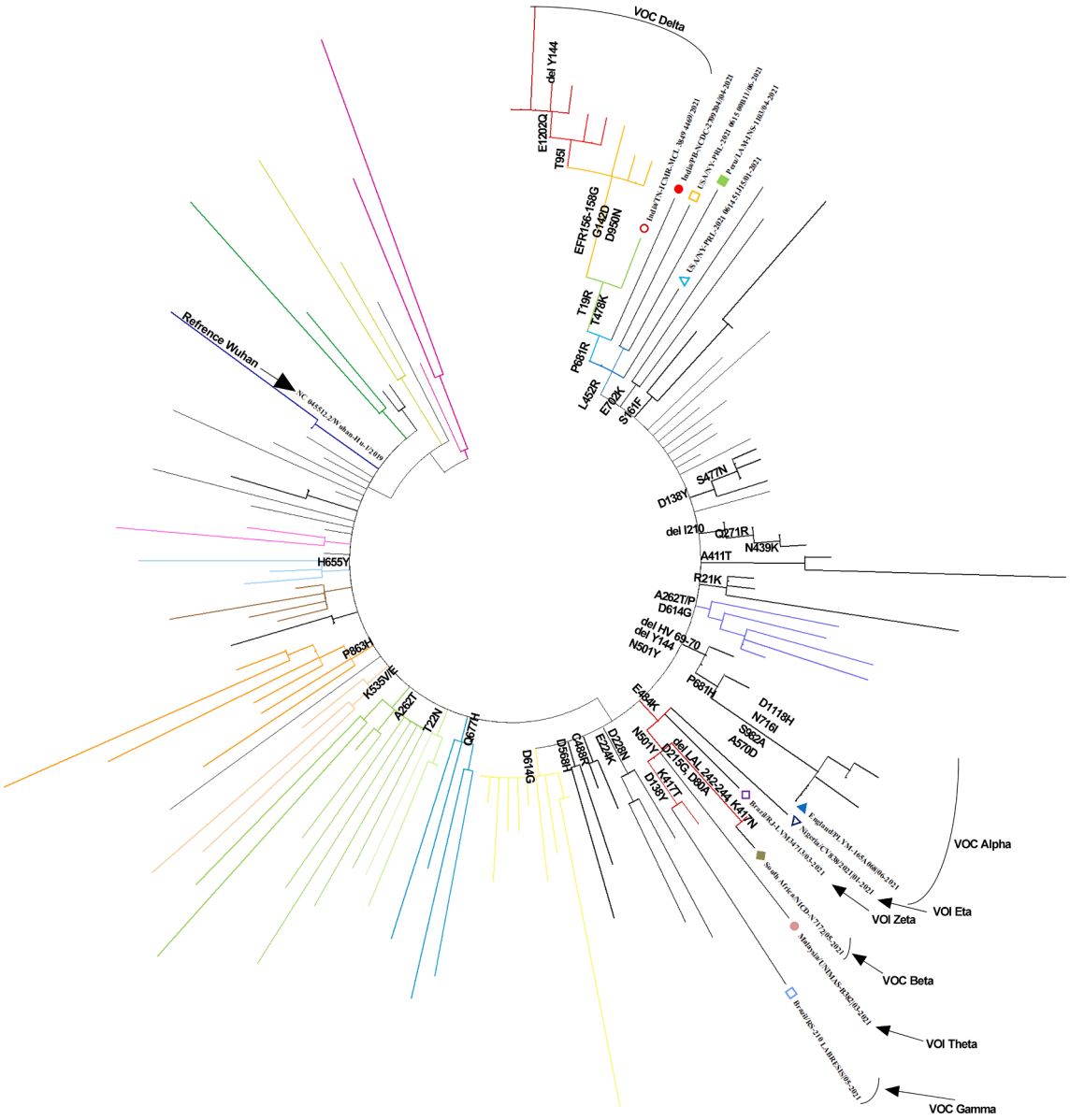

Fig. 2. Phylogenetic analysis of SARS-CoV2 strains in Iran based on spike gene. The evolutionary history was deduced through application of the Maximum Likelihood strategy and Tamura-Nei model (Tamura and Nei, 1993). The tree with the highest log likelihood (-10612.38) is demonstrated. The percentage of trees in which the associated taxa clustered together is depicted near the branches. Initial tree(s) for the heuristic search were retrieved automatically through using Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) method, and then choosing the topology with higher log likelihood amount. The tree is drawn to scale, with branch lengths measured in the quantities of substitutions per site. This analysis involves 142 nucleotide sequences. There was a total of 3822 positions in the final dataset. Evolutionary analysis was performed in MEGA X (Kumar, 2018). Representative of VOC and VOI from GISAID (11 sequences) were used for comparison and classification of Iranian SARS-COV-2 strains (131 different S gene sequences). Spike gene sequence of Wuhan-Hu-1(NC_045512.2) was used as reference. Mutation names discriminating the sequences into different branches based on their similarity are indicated on the branches and nodes. The representative VOCs and VOIs sequences are displayed with colored shapes and their names and dates of collection.

Table 7
Main characteristics of the nine emergent SARS-CoV-2 variants classified as Delta variant in Iran.

Variants	EPI_ISL_2227270	EPI_ISL_2360252	EPI_ISL_2227269	EPI_ISL_2360254	EPI_ISL_2227268	EPI_ISL_2360253	EPI_ISL_2227271	EPI_ISL_2360256	EPI_ISL_2227272
1st detection	04–2021	04–2021	04–2021	04–2021	05–2021	04–2021	04–2021	04–2021	04–2021
Detection site	Yazd	Yazd	Yazd	Yazd	Yazd	Maku	Yazd	Bushehr	Yazd
Specific mutations as indicator of a Delta variant	9 mutations: T19R, T95I, G142D, R158G, L452R, T478K, D614G, P681R, D950N 2 deletions: E156-, F157-	9 mutations T19R, T95I, G142D, R158G, L452R, T478K, D614G, P681R, D950N 2 deletions: E156-, F157-	8 mutations T19R, T95I, R158G, L452R, T478K, D614G, P681R, D950N 2 deletions: E156-, F157-	9 mutations T19R, T95I, G142D, R158G, L452R, T478K, D614G, P681R, D950N 2 deletions: E156-, F157-	9 mutations T19R, T95I, G142D, R158G, L452R, T478K, D614G, P681R, D950N 2 deletions: E156-, F157-	9 mutations T19R, T95I, G142D, R158G, L452R, T478K, D614G, P681R, D950N, 2 deletions: E156-, F157-	8 mutations T19R, G142D, R158G, L452R, T478K, D614G, P681R, D950N, 2 deletions: E156-, F157-	8 mutations T19R, G142D, R158G, L452R, T478K, D614G, P681R, D950N 2 deletions: E156-, F157-	8 mutations T19R, G142D, R158G, L452R, T478K, D614G, P681R, D950N, 2 deletions: E156-, F157-
Other mutations in S protein	1 mutation: E1202Q	4 mutations: E1202Q, C1250R, F1256L, D1257E	1 mutation: E1202Q 1 deletion: Y144-	2 mutations: L938F, E1202Q	1 mutation: P479L	1 mutation: S254F	2 mutations: D574Y, L1141W	–	1 mutation: K77T
GISAID Clade/ Pango lineage/ Varian	G/ B.1.617.2 / VOC Delta	G/ B.1.617.2 / VOC Delta	G/ B.1.617.2 / VOC Delta	G/ B.1.617.2 / VOC Delta	G/ B.1.617.2 / VOC Delta	G/ B.1.617.2 / VOC Delta	G/ B.1.617.2 / VOC Delta	G/ B.1.617.2 / VOC Delta	G/ B.1.617.2 / VOC Delta
Potential risks	• Higher transmission (Sheikh, 2021; Campbell, 2021) • Higher disease severity (Sheikh, 2021) • Escape the immune system (vaccine escape) (Sheikh, 2021; Bernal et al., 2021)								

Table 8
Main characteristics of the eleven SARS-CoV-2 variants carrying I210del mutation in Iran.

Variants	EPI_ISL_2294592	EPI_ISL_2294587	EPI_ISL_2294589	EPI_ISL_2294586	EPI_ISL_2294588	EPI_ISL_862079	EPI_ISL_862080	EPI_ISL_1014687	EPI_ISL_2455556	EPI_ISL_959277	EPI_ISL_1014684
1st detection	01–2021	12–2020	12–2020	12–2020	12–2020	11–2020	11–2020	08–2020	01–2021	08–2020	05–2020
Detection site	Shiraz	Shiraz	Shiraz	Shiraz	Shiraz	Tehran	Tehran	Tehran	Shiraz	Gilan	Urmia
Common specific mutations	D614G, I210del,	D614G, I210del,	D614G, I210del,	D614G, I210del,	D614G, I210del,	D614G, I210del,	D614G, I210del,	D614G, I210del,	D614G, I210del	D614G, I210del	I210del
Other mutations in S protein	D138Y, S477N	N439K	N439K, Q271R	N439K, N501T, Q271R	N501T, Q271R	Q314R	Q677H, D574N, D950N,	F1256L, L5F	G181V, H49Y,	–	–
GISAID Clade/Pango lineage/Varian	G/None/?	G/None/?	G/None/?	G/None/?	G/None/?	GH/ B.1.36/?	GH/ B.1.36/?	GH/ B.1.36.7/?	GH/ B.1.36/?	GH/ B.1.36/?	L/B/?

Table 9
Main characteristics of the seven emergent SARS-CoV-2 variants classified as Alpha variant in Iran.

Variants	EPI_ISL_2294590	EPI_ISL_2294591	EPI_ISL_2294585	EPI_ISL_1993554	EPI_ISL_1993552	EPI_ISL_2360250	EPI_ISL_1993555
1st detection	01–2021	01–2021	12–2020	04–2021	04–2021	05–2021	04–2021
Detection site	Shiraz	Shiraz	Shiraz	Tehran	Tehran	Kerman	Tehran
Specific mutations as indicator of a probable Alpha variant	D614G, H69del, N501Y, V70del, Y144del	D614G, H69del, N501Y, V70del, Y144del	H69del, N501Y, P681H, V70del	A570D, D614G, D1118H, H69del, L699I, N501Y, P681H, S982A, T716I, V70del, Y144del,	A570D, D614G, D1118H, H69del, N501Y, P681H, S982A, T716I, V70del, Y144del	A570D, D614G, D1118H, H69del, N501Y, P681H, S982A, T716I, V70del, V976F, Y144del	A570D, D614G, D1118H, H69del, N501Y, P681H, S982A, T716I, V70del, Y144del
Other mutations in S protein	N211del, L212I	–	–	I100T	L699I	–	–
Clade/ lineage/ Varian	G/None/?	G/None/?	O/None/?	GRY/ B.1.1.7/VOC Alpha	GRY/ B.1.1.7/VOC Alpha	GRY/ B.1.1.7/VOC Alpha	GRY/ B.1.1.7/VOC Alpha

Table 10
Main characteristics of the three SARS-CoV-2 variants carrying D138Y + S477N + D614G triple mutations in Iran.

Variants	EPI_ISL_862075	EPI_ISL_2466656	EPI_ISL_2294592
1st detection	11–2020	12–2020	01–2021
Detection site	Tehran	unknown	Shiraz
Common triple mutations	D138Y, D614G, S477N	D138Y, D614G, S477N	D138Y, D614G, S477N
Other mutations in S protein	M177I	–	I210del
GISAID Clade/ Pango lineage/ Varian	GR/ B.1.1.413/?	G/B.1/?	G/None/?
Potential risks	disrupt the epitope for mAb (Dejnirattisai, 2021) reinforce the binding of the SARS-COV-2 spike with the hACE2 receptor increases viral replication in the upper respiratory tract and enhances the (Singh, 2021) vulnerability of the virus to neutralization by antibodies (Plante, 2021)		

Table 11
Main characteristics of the two emergent SARS-CoV-2 variants classified as Beta variant in Iran.

Variants	EPI_ISL_2360255, EPI_ISL_2227273.2
1st detection	04/2021
Detection site	Hormozgan
specific mutations as indicator of a probable Beta variant	D80A, D215G, L242del, A243del, L244del, K417N, E484K, N501Y, D614G, A701V
Other mutations in S protein	–
Clade/lineage/Varian	GH/ B.1.351/ VOC Beta
Potential risks	– Higher transmission (Madhi, 2021) – Higher disease severity (Davies et al., 2021; TWR, C.A.P., 2021) – Escape the immune system (vaccine escape) (Madhi, 2021; Cele, 2021)

Table 12
Main characteristics of the possible Gamma variant in Iran.

Variant	EPI_ISL_2294576
1st detection	10-2020
Detection site	Shiraz
Specific mutations as indicator of a probable Gamma variant	D138Y, D614G, E484K, K417T, N501Y
Other mutations in S protein	S477N
Clade/lineage/Variant	G/None/?
Potential risks	– Higher transmission – Higher disease severity – - Escape the immune system (vaccine escape)

mutations of a Delta variant, also have a common E1202Q mutation in the Heptad Repeat (HR2) subdomain, in which glutamine replaces glutamic acid at position 1202. The HR1 and HR2 domains are exposed to interact with each other to form a six-helical bundle (6-HB), consequently bringing viral and cellular membranes into close proximity to permit lipid bilayer fusion (Xia, 2020). Therefore, mutations in HR1 and HR2 domains might affect heterologous 6-HBs forming, thus affecting viral/cell membrane fusion. Regarding the frequency of E1202Q mutation in Delta variant sequences from Iran, it seems that this mutation might confer an advantage to viral/cell membrane fusion process. One out of these four Delta variant sequence also showed a deletion in the NTD region of the spike protein at Y144. This Y144 deletion mutation is similar to the Y144 specific deletion mutation as indicator of the Alpha variant (Bhattarai et al., 2021), although up to the end of April 2021 this mutation has been reported in the 86 Delta variant sequences from different countries in GISAID Database. It has been shown that Y144 deletion contribute to immune escape and mostly confers resistance to most NTD-directed monoclonal antibodies (mAbs) (Suryadevara, 2021). The mutations E1202Q and Y144 deletion in combination with delta variant specific mutations might increase the concerns raised about these variants. The emergence of such local variants during the COVID-19 fifth wave in Iran needs to be further investigated for their public health impact and its possibility of becoming a new VOC. None of the Delta variant sequences from Iran (deposited in GISAID and GenBank up to the end of May 2021) carry the K417N mutation, an additional mutation in Delta variant leading to a new Delta sub-lineage referred to as Delta Plus (Roy and Roy, 2021).

We also observed an NTD deletion at position I210 in eleven sequences (6.3% of samples) sampled from May 2020 to January 2021 from four different cities. Ten out of these eleven sequences belonged to spike sequences carrying the mutation D614G (clade G); however, one sequence belonged to the Glade L, having the original amino acid aspartic acid at position 614. The NTD of the virus spike protein has appeared as a potentially mutable structure. It has deletion-prone regions that may allow the virus to escape antibody neutralization (Venkatakrishnan et al., 2021).

Among the sequences submitted from Iran, 4 sequences had all specific mutations as indicator of the Alpha variant, classified ass GRY/ B.1.1.7/VOC Alpha variant in GISAID. Two out of these four sequences each had another different mutation including I100T and L699I. Also, there were two other sequences sampled from Shiraz in January 2021, which had five specific mutations as indicator of the Alpha variant (D614G, H69del, N501Y, V70del, Y144del) and were clustered along with the Alpha variants in phylogenetic tree; however, these sequences did not contain the other specific mutations of Alpha variant including A570D, D1118H, L699I, P681H, S982A and T716I. Also, there was

Table 13
Main characteristics of the fifteen SARS-CoV-2 variants carrying A262T mutation in Iran.

Variants	*MW090851.1*	*MW055255.1*	*MW055435.1*	*MW055256.1*	*EPI_ISL_S96383*	*MW090852.1*	*EPI_ISL_S68497*	*EPI_ISL_S82025*	*EPI_ISL_S68485*	*EPI_ISL_S68503*	*EPI_ISL_S91324*	*MW090849.1*	*MW090862.1*	*MW055257.1*	*EPI_ISL_9122O6*
1st detection	03–2020	05–2020	04–2020	05–2020	04–2020	03–2020	03–2020	08–2020	03–2020	04–2020	04–2020	03–2020	04–2020	05–2020	2020–08
Common specific mutations	A262T	A262T	A262T	A262T	A262T	A262T	A262T	A262T	A262T	A262T	A262T	A262T	A262T	A262T	A262T
Other mutations in S protein	T302A D614G R21K, R44K, D111N, S112P	N542T, V620F, K77N D614G	S459Y, D614G	N487I, V620I, P225T, D614G	D614G, R328T	N536K, T791P, K811N, T827A	D843E, N487H, P1263L T874P, V987F, Y144D, Y200N	A672P, C617R, D287Y, D586E, E583A, E661K, G652R, N603K, N606Y, T602K, V289E	A1015, A1026, D843G, K835N, L841I, P863H, S1021P	K150I, P863H, Q675R, V781F, Y873N	D467N, Q675R	Q498Y, P499T, T716P, V11I, L176F, L179P I726F	N487I, N641K V11I, T22N, Q23H, H66L, N74H,	S758I, P863H, N925Y, G932A, D614G Q115H, V171L, K182N, T941P	D614G, P412S, P863H, V407L

another sequence sampled from Shiraz in Dec 2020 which had the original amino acid aspartic acid at position 614 and contained four specific mutations as indicator of the Alpha variant and was clustered along with the Alpha variants in phylogenetic tree.

We also observed triple co-occurring mutations, namely D138Y (in the NTD), S477N (in the RBD) with the well-established D614G mutation (near the furin cleavage site) in three spike sequences from Tehran and Shiraz sampled from Nov 2020 to Jan 2021. The mutations D138Y and S477N are the same as those found in the NTD of VOC Gamma (P.1) and RBD of VOI of lota, respectively. It has been shown that these two mutations could contribute to reduced neutralization by some mAbs, convalescent plasma, and sera from vaccines. Moreover, they have higher binding affinity for human ACE2 (Dejnirattisai, 2021; Singh, 2021). These co-occurring mutations might increase the concerns raised about these variants and needs to be further investigated for their public health impact and the possibility of becoming a VOC.

We also detected two complete SARS-COV2 genome sequences sampled from Hormozgan in Apr 2021 (with identical spike protein sequences) carrying all the specific mutations as indicator of a Beta variants, and classified as VOC Beta GH/501Y.V2 (B.1.351 + B.1.351.2 + B.1.351.3) in GISAID. There are particular concerns attributed to this variant, including increased transmission (Pearson, 2021), significantly reduced susceptibility to the mAb treatment and reduced neutralization by convalescent and post-vaccination sera (Madhi, 2021; Cele, 2021). However, there are no other sequences of VOC Beta variants from Iran, deposited in GISAID after Apr 2021 until Aug 2021.

There was another sequence sampled from Shiraz in Oct 2020 that contained five specific mutations (D138Y, D614G, E484K, K417T, N501Y) as indicator of the Gamma variant (GR/501Y.V3 (P.1 + P.1.x)) and was clustered along with the representative Gamma variant spike sequence in phylogenetic tree. This sequence also had an additional mutation S477N in RBD. The P.1 variant also has been raised particular concerns, including increased transmissibility, possible increased risk of hospitalization, moderate reduction in neutralizing activity of mAb and convalescent and post-vaccination sera (US Food and Drug Administration; Wang, 2021). However, there are no other sequences of VOC Gamma variants from Iran, deposited in GISAID after Oct 2021 until Aug 2021.

We also found that the A262T mutation is the most frequent mutation (15 sequences; 8.5% of samples) among the Iranian SARS-COV2 spike sequences sampled from Mar 2020 to Aug 2020; however, this mutation has been disappeared after Aug 2020, probably due to the emergence of new variants. We found the mutation A262T in both spike sequence with the original D614 (8 out of 15) and with the well-established D614G mutation (7 out of 15). This mutation has been deposited in GISAID database in Mar 2020 for the first time from three countries including Iran, USA and Australia. Currently, the A262T mutation is seen in some variants from different countries with a very low frequency (221 SARS-COV2 sequences in GISAID).

We also assessed the distribution of SARS-COV-2 clades in Iran. This analysis determined the sequential prevalence of O, GH, GRY and G clades, the last one being characterized by D614G variant. We also suggested that currently more than 95% of the Iranian strains belong to the two major branches of the G clade, i.e., GR and GH clades and the newly introduced GRY clade. It is worth mentioning that differences in virulence of clades might at least partially explain differences in mortality rates or rate of transmission among different countries (Brufsky, 2020). Consistent with this speculation, the comprehensive sequence analyses in each region have facilitated understanding the specific geographic distribution of SARS-CoV-2 variants. This information has practical significance in defining clinical and political strategies for management of this disorder in each geographic region.

Cumulatively, we reported the detected mutations in spike protein in Iranian SARS-CoV-2 samples. We found that the four current VOCs – Alpha, Beta, Gamma and Delta – are circulated in Iran. The Delta variant is notably more transmissible than other variants, and is expected to

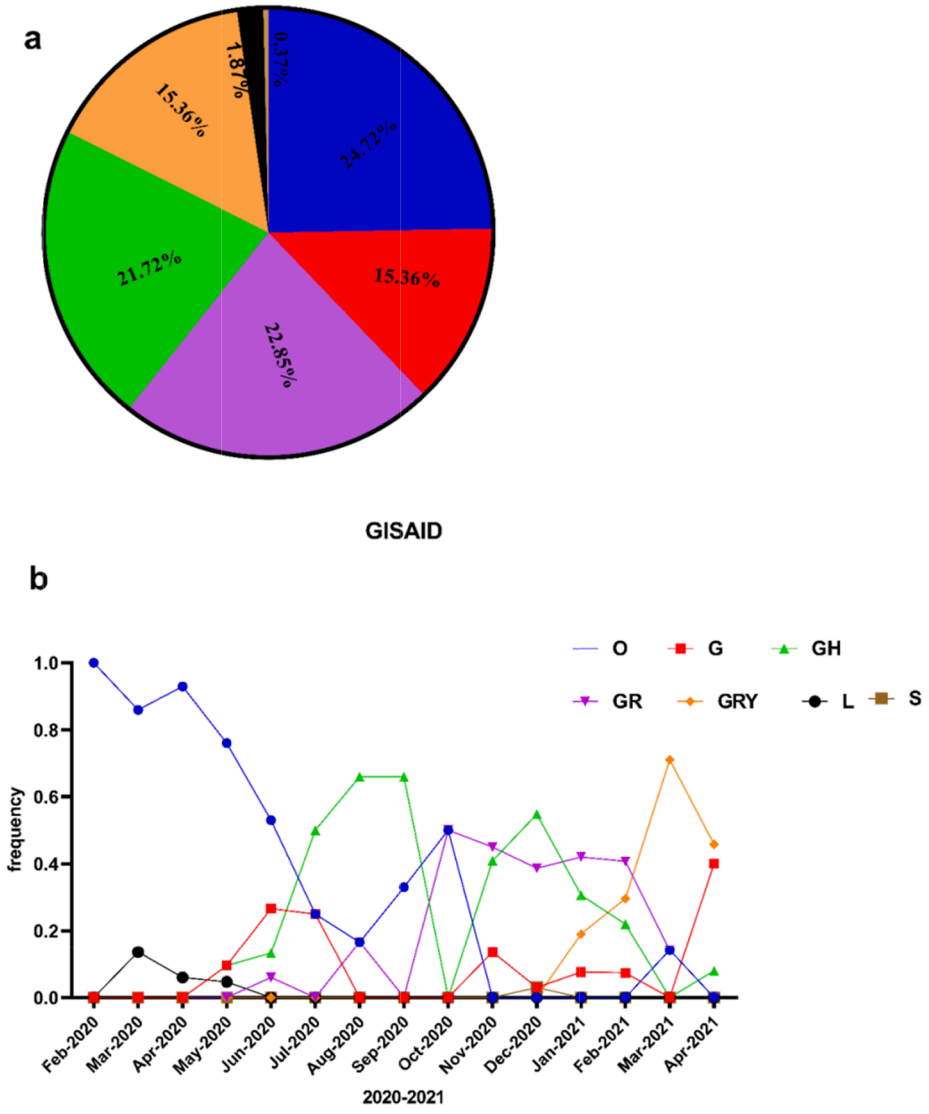

Fig. 3. The distribution of Iranian SARS-CoV-2 clades over time.

become a dominant variant. However, some of the Delta variants in Iran carry an additional mutation, E1202Q in the HR2 subdomain that might confer an advantage to viral/cell membrane fusion process. The mutations reported in the current study have importance in the development of antibody, vaccine, and drugs as well as designing management strategies for this disorder.

CRediT authorship contribution statement

Solat Eslami: Conceptualization, Methodology, Software. **Mark C. Glassy:** Reviewing and Editting. **Soudeh Ghafouri-Fard:** Writing – original draft.

Declaration of Competing Interest

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

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2021.146113.

References

Bernal, J.L., et al., 2021. Effectiveness of COVID-19 vaccines against the B. 1.617. 2 variant. medRxiv.
Bhattarai, N., et al., 2021. Structural and dynamical differences in the spike protein RBD in the SARS-CoV-2 Variants B. 1.1. 7 and B. 1.351. J. Phys. Chem. B.
Braun, J., et al., 2020. Presence of SARS-CoV-2 reactive T cells in COVID-19 patients and healthy donors. MedRxiv.
Braun, K.M., et al., 2021. Transmission of SARS-CoV-2 in domestic cats imposes a narrow bottleneck. PLoS Pathog. 17 (2), e1009373.
Brufsky, A., 2020. Distinct viral clades of SARS-CoV-2: Implications for modeling of viral spread. J. Med.Virol. 92 (9), 1386–1390.
Campbell, F., et al., 2021. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. Eurosurveillance 26 (24), 2100509.
Cao, Y., et al., 2020. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. Cell Discovery 6 (1), 1–4.
Cavallo, L., Oliva, R., 2020. D936Y and other mutations in the fusion core of the SARS-CoV-2 spike protein heptad repeat 1 undermine the post-fusion assembly.
Cele, S., et al., 2021. Escape of SARS-CoV-2 501Y. V2 from neutralization by convalescent plasma. Nature 593 (7857), 142–146.
Chaudhari, A., et al., 2021. E156G and Arg158, Phe-157/del mutation in NTD of spike protein in B. 1.617. 2 lineage of SARS-CoV-2 leads to immune evasion through antibody escape. bioRxiv.
Coutard, B., et al., 2020. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antiviral Res. 176, 104742.
Davies, N.G., et al., 2021. Increased mortality in community-tested cases of SARS-CoV-2 lineage B. 1.1.7. Nature 593 (7858), 270–274.
Dejnirattisai, W., et al., 2021. Antibody evasion by the P. 1 strain of SARS-CoV-2. Cell 184 (11), 2939–2954 e9.

Di Giacomo, S., et al., 2021. Preliminary report on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Spike mutation T478K. J. Med.Virol.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32 (5), 1792–1797.
Flemming, A., 2021. SARS-CoV-2 variant evades antibodies whilst maintaining fitness. Nature Rev. Immunol. 21 (3), 136-136.
Grabowski, F., et al., 2021. SARS-CoV-2 variant of concern 202012/01 has about twofold replicative advantage and acquires concerning mutations. Viruses 13 (3), 392.
Guo, E., Guo, H., 2020. CD8 T cell epitope generation toward the continually mutating SARS-CoV-2 spike protein in genetically diverse human population: implications for disease control and prevention. PLoS ONE 15 (12), e0239566.
Guruprasad, L., 2021. Human SARS CoV-2 spike protein mutations. Proteins 89 (5), 569–576.
Harvey, W.T., et al., 2021. SARS-CoV-2 variants, spike mutations and immune escape. Nat. Rev. Microbiol. 19 (7), 409–424.
Jangra, S., et al., 2021. SARS-CoV-2 spike E484K mutation reduces antibody neutralisation. Lancet Microbe.
Kemp, S.A., et al., 2021. SARS-CoV-2 evolution during treatment of chronic infection. Nature 592 (7853), 277–282.
Kemp, S.A., et al., 2021. Recurrent emergence and transmission of a SARS-CoV-2 Spike deletion H69/V70.
Khan, A., et al., 2021. Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data. J. Cell. Physiol.
Kumar, S., et al., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35 (6), 1547.
Laha, S., et al., 2020. Characterizations of SARS-CoV-2 mutational profile, spike protein stability and viral transmission. Infect. Gen. Evolut. 85, 104445.
Lan, J., et al., 2020. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 1–6.
Lazarevic, I., et al., 2021. Immune evasion of SARS-CoV-2 emerging variants: what have we learnt so far? Viruses 13 (7), 1192.
Li, Q., et al., 2020. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. Cell 182 (5), 1284–1294 e9.
Liu, C., et al., 2021. Reduced neutralization of SARS-CoV-2 B. 1.617 by vaccine and convalescent serum. Cell.
Liu, Z., et al., 2020. Landscape analysis of escape variants identifies SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. Available at SSRN 3725763.
Luan, B., Wang, H., Huynh, T., 2021. Enhanced binding of the N501Y-mutated SARS-CoV-2 spike protein to the human ACE2 receptor: insights from molecular dynamics simulations. FEBS Lett. 595 (10), 1454–1461.
Lubinski, B., et al., 2021. Functional evaluation of proteolytic activation for the SARS-CoV-2 variant B. 1.1. 7: role of the P681H mutation. bioRxiv.
Madhi, S.A., et al., 2021. Efficacy of the ChAdOx1 nCoV-19 Covid-19 vaccine against the B. 1.351 variant. N. Engl. J. Med. 384 (20), 1885–1898.
Mishra, A., et al., 2021. Natural infection of SARS-CoV-2 delta variant in Asiatic lions (Panthera leo persica) in India. bioRxiv.
Motozono, C., et al., 2021. SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. Cell Host Microbe.
Pearson, C., et al., 2021. Estimates of severity and transmissibility of novel South Africa SARS-CoV-2 variant 501Y. V2.
Plante, J.A., et al., 2021. Spike mutation D614G alters SARS-CoV-2 fitness. Nature 592 (7852), 116–121.
Roy, B., Roy, H., 2021. The Delta Plus variant of COVID-19: Will it be the worst nightmare in the SARS-CoV-2 pandemic? J. Biomed. Sci. 8 (1), 1–2.
Saito, A., et al., 2021. SARS-CoV-2 spike P681R mutation enhances and accelerates viral fusion. bioRxiv.
Schrörs, B., et al., 2021. Large-scale analysis of SARS-CoV-2 spike-glycoprotein mutants demonstrates the need for continuous screening of virus isolates. PLoS ONE 16 (9), e0249254.
Sheikh, A., et al., 2021. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. The Lancet.
Singh, A., et al., 2021. Serine 477 plays a crucial role in the interaction of the SARS-CoV-2 spike protein with the human receptor ACE2. Sci. Rep. 11 (1), 1–11.
Starr, T.N., et al., 2021. Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. Cell Rep. Med. 2 (4), 100255.
Sun, C., et al., 2020. SARS-CoV-2 and SARS-CoV spike-RBD structure and receptor binding comparison and potential implications on neutralizing antibody and vaccine development. Biorxiv.
Suryadevara, N., et al., 2021. Neutralizing and protective human monoclonal antibodies recognizing the N-terminal domain of the SARS-CoV-2 spike protein. Cell 184 (9), 2316–2331 e15.
Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10 (3), 512–526.
Tomkins-Tinch, C.H., et al., Detection of the recurrent substitution Q677H in the spike protein of SARS-CoV-2 in cases descended from the lineage B. 1.429.
TWR, C.A.P., et al., 2021. Estimates of severity and transmissibility of novel South Africa SARS-CoV-2 variant501Y. V2. Centre for Mathematical Modelling of Infectious Diseases.
US Food and Drug Administration. Fact sheet for health care providers emergency use authorization (EUA) of bamlanivimab and etesevimab. Available at: https://www.fda.gov/media/145802/download.
Venkatakrishnan, A., et al., 2021. Antigenic minimalism of SARS-CoV-2 is linked to surges in COVID-19 community transmission and vaccine breakthrough infections. medRxiv.
Wang, N., et al., 2020. Subunit vaccines against emerging pathogenic human coronaviruses. Front. Microbiol. 11, 298.
Wang, R., et al., 2021. Vaccine-escape and fast-growing mutations in the United Kingdom, the United States, Singapore, Spain, India, and other COVID-19-devastated countries. Genomics 113 (4), 2158–2170.
Wang, P., et al., 2021. Increased resistance of SARS-CoV-2 variant P. 1 to antibody neutralization. Cell host & microbe 29 (5), 747–751 e4.
West, Jr., A.P., et al., 2021. Detection and characterization of the SARS-CoV-2 lineage B. 1.526 in New York. bioRxiv.
Wibmer, C.K., et al., 2021. SARS-CoV-2 501Y. V2 escapes neutralization by South African COVID-19 donor plasma. Nat. Med. 27 (4), 622–625.
Xia, S., et al., 2020. Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. Cell. Mol. Immunol. 17 (7), 765–767.
Yao, X., et al., 2020. In vitro antiviral activity and projection of optimized dosing design of hydroxychloroquine for the treatment of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Clin. Infect. Dis.
Zhang, L., et al., 2020. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. Nat. Commun. 11 (1), 1–9.
Zhang, X., et al., 2021. Inhibitor screening of Spike variants reveals the heterogeneity of neutralizing antibodies to COVID-19 infection and vaccination. medRxiv.
Zhou, Z.-J., Qiu, Y., Ge, X.-Y., 2020. Characterization of the substitution hotspots in SARS-CoV-2 genome using BioAider and detection of a SR-rich region in N protein providing further evidence of its animal origin. bioRxiv.