



Identification of novel mutations in SARS-COV-2 isolates from Turkey

Shazia Rehman^{1,2} · Tariq Mahmood³ · Ejaz Aziz⁴ · Riffat Batool⁵

Received: 24 July 2020 / Accepted: 1 September 2020 / Published online: 10 October 2020
© Springer-Verlag GmbH Austria, part of Springer Nature 2020

Abstract

SARS-CoV-2 was first detected in the city of Wuhan, Hubei Province, China. In this study, we identified 11 unique mutations in viral SARS-COV-2 isolates from Turkey. Nine of them cause structural alterations in the S protein, nsp2, nsp3, nsp4 and nsp12 regions. The mutations identified here might have significant functional implications that need to be addressed in future studies in the context of vaccine engineering and therapeutic interventions. Moreover, transmission and phylogenetic analysis revealed multiple independent sources of introductions of SARS-CoV-2 into Turkey and a close relationship to the isolates from Saudi Arabia.

The first case of the novel coronavirus outbreak in humans was reported in Wuhan, China. The disease was named COVID-19 by WHO, and the virus was named SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) by the International Committee on Taxonomy of Viruses [1]. Since the emergence of COVID-19, based on the WHO report of July 8, 2020, more than 11,591,595 confirmed cases were reported in 147 countries, with 537,859 deaths, due to rapidly spreading SARS-CoV-2 [2]. Genetically different coronaviruses are spread in birds, humans and other mammals and can lead to severe diseases of the intestine, liver, nervous system, and respiratory system. Turkey straddles eastern Europe and western Asia and is a major travel hub. According to the World Health Organization's report

dated April 24, 2020, Turkey ranks sixth in the European region in terms of COVID-19 disease, after Spain, Italy, Germany, the United Kingdom, and France [3]. The first case of COVID-19 in Turkey was reported on March 11, 2020, and two months after the first case, on May 11, 2020, the Turkish Ministry of Health declared that the number of COVID-19 cases had reached 139,771, with 3841 deaths [4]. In the light of the coronavirus outbreak, the present study was designed with the aim to characterize notable genetic features of SARS-CoV-2 from Turkey and to identify some novel mutations in the spike protein (S), nucleocapsid protein (N), and non-structural proteins (nsp2, nsp3, nsp4, nsp6, nsp12/RdRP). Furthermore, transmission and phylogenetic analysis were also conducted to provide significant insight into the spread of the virus within Turkey.

For analysis, a total of 80 genome sequences of virulent strains from Turkey that had been uploaded to the NCBI (<https://www.ncbi.nlm.nih.gov/genbank>) and GISAID (<https://www.gisaid.org/>) databases as of May 4, 2020, were retrieved and compared to genome sequences from Saudi Arabia, Iran, America, China, Pakistan, Denmark, Spain, and Italy (Table S1). These sequences were first aligned using the Clustal W program, and a phylogenetic tree was constructed based on this alignment using the maximum-likelihood method in MEGA X software with 1000 replicates [5]. For identification of mutations, sequences of SARS-CoV-2 isolates were compared with a reference sequence (MN908947.1) from Wuhan. The secondary structure of the S protein and nonstructural proteins of SARS-COV-2 was predicted using CFSSP (Chou and Fasman secondary structure prediction), an online server [6].

Communicated by Tim Skern.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00705-020-04830-0>) contains supplementary material, which is available to authorized users.

✉ Shazia Rehman
shaziarehman7@gmail.com

- ¹ Department of Botany, Rawalpindi Women University, Rawalpindi, Pakistan
- ² Department of Botany, Govt. Gordon College Rawalpindi, Rawalpindi, Pakistan
- ³ Department of Plant Sciences, Quaid-i-Azam University, Islamabad 45320, Pakistan
- ⁴ Department of Botany, GDC Khanpur, Haripur, Pakistan
- ⁵ University Institute of Biochemistry and Biotechnology, PMAS-UAAR, Rawalpindi, Pakistan

Based on mutation analysis, 59 out of 80 isolates from Turkey contained a signature 23,403A > G (D614G) mutation in the spike glycoprotein (S), which is clearly indicative of a very frequent mutation (73%). Most samples with the D614G mutation were strongly associated with two other mutations (3037 C > T and 14,408C > T) in ORF1ab region (Table 1). These co-occurring mutations have been described recently as a characteristic of one of the major SARS-CoV-2 variants occurring in Europe. Regarding the ORF 1ab region, we also identified previously reported single-nucleotide polymorphisms (SNPs) at positions 14,408 (C > T), 3037 (C > T), 11,083 (G > T), 1397 (G > A), 18,877 (C > T), 1059 (T > A) and 8782 (C > T). The 14,408C > T (P4715L) and 3037 C > T (F106F) variants in ORF1ab were found to occur at high frequency and are presumed to be linked, causing mutations in the RNA-dependent RNA polymerase (RdRP/nsp12) and nsp3 gene, respectively. RdRP/nsp12 is a key component of the replication/transcription machinery, and therefore, the leucine mutation at position 4715 of RdRP/nsp12 may potentially affect its function, possibly increasing the viral mutation rate. Moreover, the proline-to-leucine mutation was consistently observed in previous reports as a frequent mutation in Europe (51.6%) and North America (58.1%) [7, 8]. The variation C3037T has been reported to cause a synonymous mutation in the region encoding nsp3 and was seen in 57 isolates (71%) from Turkey. Consistent with other studies, C3037T, C14408T and A23403G were the most common mutations (73%) that were found together in isolates from Turkey.

Other key variations observed in the present report includes 25,563G > T (Q57H) in ORF3a, along with a consecutive series of three variations at positions 28,881 (G > A), 28,882 (G > A) and 28,883 (G > C) in the N protein. The triple mutation 28,881–28,883, which results in a change of two amino acids, 203–204:RG > KR, is known to play a critical role in virion assembly and structure and had been found frequently in US strains [9]. In present study, this tri-nucleotide mutation in the N protein was observed in eight samples from Turkey, including one from a patient (EPI_ISL_429870) with a history of travel to Saudi Arabia. This mutation was accompanied by the mutations 241:C > T, 3037:C > T and 14,408:C > T. The missense mutation G11083T, conferring an amino acid change from leucine (L) to phenylalanine (F) in non-structural protein 6 (nsp6) at position 3606, was present in 25 samples. It has been observed in Spain, Italy, and Iran, as well in this study (Table S2). Previously, it was reported as an infrequent mutation from Japan, the Netherlands, and Australia [10, 11]. Another substitution (G > A) was found within the region of ORF1a encoding nsp2 at position 1397 and was seen in 26% (21/80) of the isolates, resulting in an amino acid change from valine to isoleucine (V378I) that would not

affect the isoelectric point. In a study by Pachetti et al. [8], the V378I substitution was mainly observed in isolates from Oceania and less frequently in Asia and North America.

Our secondary structure prediction analysis also highlighted 11 unique mutations in viral SARS-COV-2 isolates from Turkey in the spike (S) protein and non-structural proteins (nsp2, nsp3, nsp4, and nsp12/RdRP) (Table 1). Of these 11 mutations, nine of them have been found to be involved in structural alterations at different sites. Three of them (A771V, T1238I and G1251V) cause alterations in the structure of the S protein, while the rest of them induce structural changes in nsp2 (A206T, R207C, T265I), nsp3 (A1824V), nsp4 (M2796I) and nsp12 (A4489V).

Among novel mutants in the S protein, a substitution mutation (I468V) at position 468 was found in four Turkish isolates (EPI_ISL_437316, EPI_ISL_437322, EPI_ISL_437323 and EPI_ISL_437324) in the receptor-binding domain (RBD) of the spike protein that was not reported previously in strains from any other country. At this site, isoleucine (I) was replaced by valine (V) at position 468. Since both amino acids are hydrophobic in nature, having C beta branched residues, this mutation is unlikely to cause a functional change in the protein, as predicted by our secondary structure analysis (Fig. 1a). Nevertheless, if this site is prone to mutate further, it is possible that such mutations could affect the binding of the S protein to ACE2 receptor. Similarly, two strains (EPI_ISL_437314, EPI_ISL_437316) contained a missense mutation at position 771 exchanging alanine (A) for valine (V) in the S protein. Secondary structure predictions demonstrated that this could disrupt the helix structure, favoring a β -sheet at positions 770 and 771 due to addition of (Fig. 1b). The alanine-to-valine substitution as residue 771 was previously seen in one Belgian strain [12].

Additionally, a threonine (polar amino acid)-to-isoleucine (non-polar amino acid) substitution (T1238I) and a change of glycine to a bulkier valine at position 1251 (G1251V) were specifically found in strains EPI_ISL_429870 and EPI_ISL_437317, respectively. A secondary structure prediction revealed that a change of threonine to isoleucine makes this position hydrophobic and favors the formation of an additional helix at position 1238 (Fig. 1c), while the G1251V mutation favors the formation of six additional sheets from position 1448 to 1253 and a helix at site 1254 (Fig. 1d). Earlier reports have suggested that substitution mutations of threonine to isoleucine and alanine to valine in the Ebola virus (EBOV) glycoprotein (GP) increase the infectivity in humans [13, 14]. Therefore, it is possible that the mutant residues at positions 771, 1238 and 1254 in the spike protein of EBOV may lead to an alteration in the way the spike interacts with the receptor, changing the infectivity of the virus, as these mutations lie in the S2 subunit of the S protein.

Table 1 Mutations identified in 80 SARS-COV-2 genome sequences from Turkey. Unique mutations in the S protein, nssp2, nsp3, nsp4, nsp12 are indicated by *

Mutations	Amino acid change	Gene	Type of mutation	Number of isolates
23,403 A>G	D614G	S	Missense	59
14,408 C>T	P4715L	ORF1ab	Missense	58
3037C>T		ORF1ab	Synonymous	57
241 C>T		5'UTR	Non coding	42
25563G>T	Q57H	ORF3a	Missense	31
18877C>T	L6205L/L280L	ORF1ab	Synonymous	29
11083G>T	L3606F	ORF1ab	Missense	25
28688 T>C		N	Synonymous	23
28,881 G>A	R203K	N	Missense	23
28882G>A		N	Synonymous	23
28,883 G>C	G204R	N	Missense	23
29742G>T		3'UTR	Non coding	22
1397G>A	V378I	Orf1ab	Missense	21
C26549T		M	Synonymous	13
26735C>T		M	Synonymous	13
884C>A	R207S	Orf1ab/nsp2	Missense	11*
8653G>T	M2796I	Orf1ab/nsp4	Missense	11*
19839 T>C		Orf1ab	Synonymous	11
7765C>A		Orf1ab/nsp3	Synonymous	09*
17690C>T	S5809L	Orf1ab	Missense	08
228C>T		5'UTR		08
28854C>T	S194L	N	Missense	08*
22444C>T	D294D	S	Synonymous	08*
2113C>T		Orf1ab/nsp2	Synonymous	07*
12809C>T	L4182F	ORF1ab	Missense	04
10702C>T		Orf1ab	Synonymous	04
26551 T>C	V10A	M	Missense	04
22964A>T	I468V	S	Missense	04*
5736 C>T	A1824V	Orf1ab/nsp3	Missense	03*
13476C>T		ORF1ab	Synonymous	03*
3903C>T	P1213L	ORF1ab/nsp3	Missense	03*
881G>A	A206T	ORF1ab/nsp2	Missense	03*
25611C>A		ORF3a	Synonymous	03
8782 C>T		Orf1ab	Synonymous	02*
1059C>T	T265I	ORF1ab	Missense	02*
23874C>T	A771V	S	Missense	02*
25275C>T	T1238I	S	Missense	01*
23929C>T	Y789Y	S	Synonymous	01*
22468G>A	T302T	S	Synonymous	01*
25314G>T	G1251V	S	Missense	01*
2416C>T	Y717Y	ORF1ab/nsp2	Synonymous	01*
13730C>T	A4489V	ORF1ab	Missense	01
28144 T>C	L84S	ORF8	Missense	01
5182 T>C		ORF1ab	Synonymous	03
27,103>T	A194V	M	Missense	01
5477 C>T	H1738Y	ORF1ab	Missense	01
6402C>T	P2046L	ORF1ab	Missense	01
12741C>A	T4159K	ORF1ab/nsp8	Missense	01
2997C>T	S911F	ORF1ab	Missense	01

Table 1 (continued)

Mutations	Amino acid change	Gene	Type of mutation	Number of isolates
13376A>G	T4371A	ORF1ab	Missense	01
19484C>T	A6407V	ORF1ab	Missense	01
1437C>T	S391F	ORF1ab	Missense	01
C7834T		ORF1ab	Synonymous	01
944G>A	G227S	ORF1ab	Missense	01
26735C>T		M	Synonymous	01
14178C>T		ORF1ab	Synonymous	01
25549C>T	L53F	ORF3a	Missense	01
16247C>T	A5328V	ORF1ab	Missense	01
6078C>T	A1938V	ORF1ab	Missense	01
16616C>T	T5451I	ORF1ab	Missense	01
27354A>G		ORF6	Synonymous	01
20464G>A	D6734N	ORF1a	Missense	01
9667 T>C		ORF1ab	Synonymous	01
29197C>T		N	Synonymous	01
7303C>T		ORF1ab	Synonymous	01
3464C>T	H1067Y	ORF1ab	Missense	01
5192C>T		ORF1ab	Synonymous	01
20629C>T	H6789Y	ORF1ab	Missense	01
20799 T>C		ORF1ab	Synonymous	01
28391C>T	R40C	N	Missense	01
25611C>A		ORF3a	Synonymous	03
20668G>A	A6802T	ORF1ab	Missense	01
27476C>T	T28I	ORF7a	Missense	01

Among the synonymous SNPs in the S protein, the C23929T (Y789Y) mutation was observed in only one Turkish sequence (EPI_ISL_455719) in the S gene. It was reported previously with high prevalence in Indian strains (39.13%) and in one US strain (EPI_ISL_436,898) [7, 12], but not in any other European strains so far. C22444T (S protein) and C28854T substitution (N protein) were observed to be uncommon co-occurring mutations in viral isolates from Turkey (8 samples) and Saudi Arabia (9 samples) (Table S2), which may provide evidence for a travel-associated origin of these mutations.

The nsp2 protein is postulated to play an important role in the host cell survival pathway via its interaction with prohibitin (PHB) and prohibitin 2 (PHB2) [15]. The A206T, R207C and T265I mutations were predicted to cause structural alterations within the nsp2 domain (Fig. 1e, f and g). A change of a nonpolar amino acid (alanine) to a polar amino acid (threonine) at position 206 (G881A mutation) in the nsp2 protein resulted in the predicted loss of an α -helix at positions 203, 204 and 206 with the addition of a β -sheet at positions 203 and 204 (Fig. 1e), and this mutation appeared only in three sequences (EPI_ISL_480239, EPI_ISL_428723, EPI_ISL_429863) from Turkey. Similarly, an 884C>T mutation, which resulted in a change of

arginine (R) to cysteine (C) at position 207 was present in 11 Turkish isolates and one Pakistani isolate (MT240479.1) (Table S2). The R207C mutation resulted in the replacement of an α -helix with a sheet structure at position 204 and an additional turn at position 210 (Fig. 1f). Furthermore, two Turkish samples (EPI_ISL_437309; EPI_ISL_437315), and one strain each from Spain (EPI_ISL_428688) and Denmark (EPI_ISL_437668) had a T265I mutation, which resulted in the addition of sheet structure at position 266 when threonine is substituted by isoleucine at position 265 (Fig. 1g). Since all of these mutations were identified in the nsp2 domain, possibly causing structural alterations, it is essential to consider the mutational spectrum when designing new antiviral therapeutics targeting the viral ORF1ab. In addition, a synonymous mutation (2416C>T) within nsp2 was found in two infected individuals (EPI_ISL_428712; EPI_ISL_437332) from Turkey with a history of travel from Iran, and in three samples (EPI_ISL_468162; EPI_ISL_468161;

Fig. 1 Prediction of secondary structure in the S protein, nsp2, nsp3, nsp4, nsp12 regions. **a–d** Mutations in the S protein. **e–g** Mutations in nsp2; **h–i** Mutations in nsp3; **j** Mutations in nsp4; **k** Mutations in nsp12. Small rectangular boxes indicate mutated residues Differences in secondary structure between Wuhan and Turkish isolates are indicated by black boxes

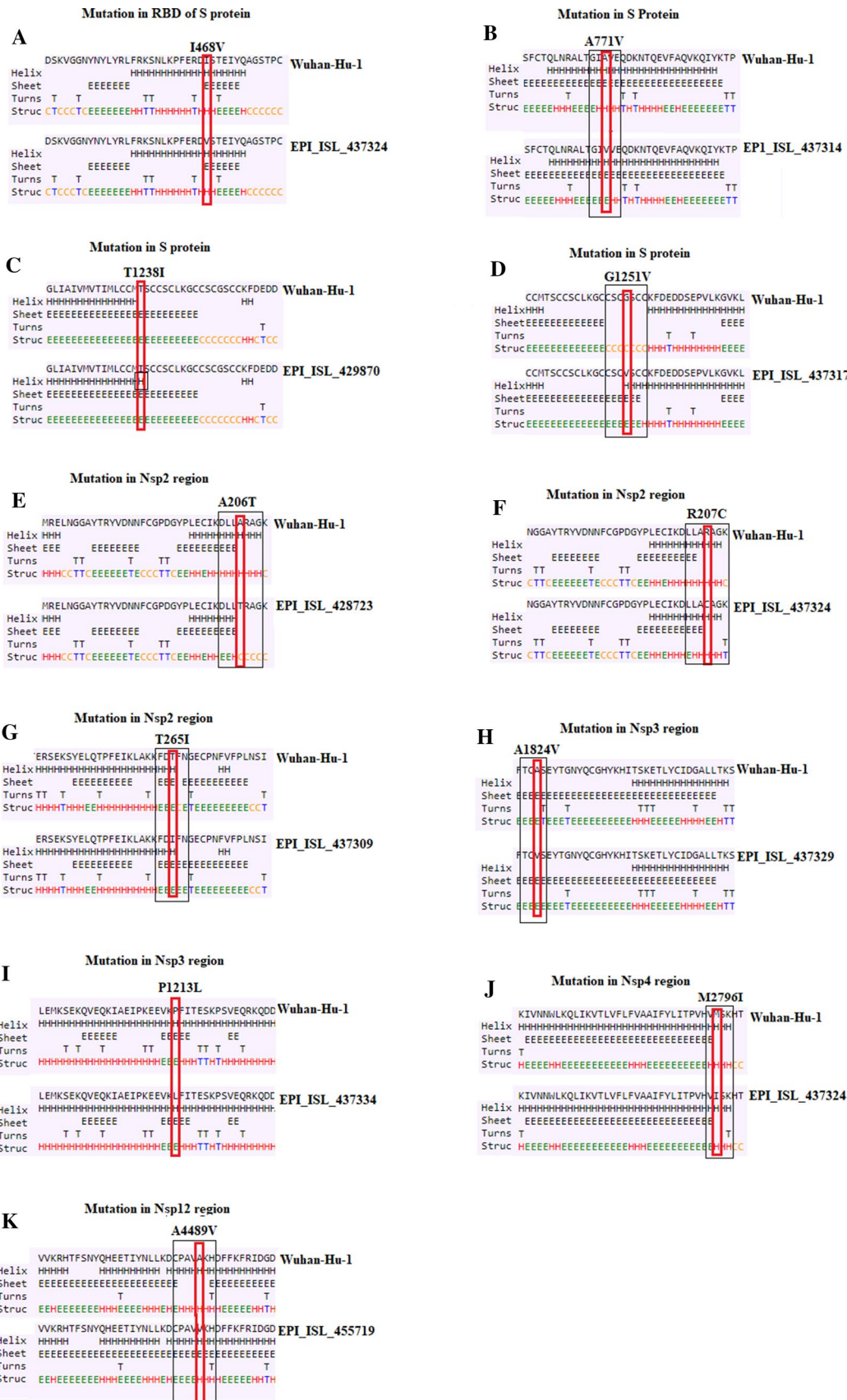


Table 2 List of Turkish strains from individuals with a history of travel

Accession no	Virus name	Location	Collection date	Travel history
EPI_ISL_429866	hCoV-19/Turkey/HSGM-4236/2020	Turkey / Afyon/west	2020-03-16	Saudi Arabia
EPI_ISL_429869	hCoV-19/Turkey/HSGM-4701/2020	Turkey / Konya	2020-03-17	Saudi Arabia
EPI_ISL_428716	hCoV-19/Turkey/HSGM-5711/2020	Turkey / Ankara	2020-03-18	Saudi Arabia
EPI_ISL_428714	hCoV-19/Turkey/HSGM-5602/2020	Turkey / Kastamonu	2020-03-18	Saudi Arabia
EPI_ISL_429862	hCoV-19/Turkey/HSGM-8970/2020	Turkey / Ankara	2020-03-22	Saudi Arabia
EPI_ISL_437332	hCoV-19/Turkey/HSGM-4698/2020	Turkey / Istanbul	2020-03-18	Iran
EPI_ISL_428718	hCoV-19/Turkey/HSGM-5770/2020	Turkey / Kocaeli	2020-03-19	Saudi Arabia
EPI_ISL_437328	hCoV-19/Turkey/HSGM-7668/2020	Turkey / Tekirdag	2020-03-19	Saudi Arabia
EPI_ISL_437326	hCoV-19/Turkey/HSGM-1432/2020	Turkey / Istanbul	2020-03-19	Iran
EPI_ISL_437325	hCoV-19/Turkey/HSGM-1495/2020	Turkey / Istanbul	2020-03-19	Iran
EPI_ISL_437324	hCoV-19/Turkey/HSGM-1476/2020	Turkey / Istanbul	2020-03-19	Iran
EPI_ISL_437327	hCoV-19/Turkey/HSGM-1458/2020	Turkey / Agri	2020-03-19	Iran
EPI_ISL_437323	hCoV-19/Turkey/HSGM-1481/2020	Turkey / Istanbul	2020-03-19	Taiwan
EPI_ISL_437319	hCoV-19/Turkey/HSGM-1490/2020	Turkey / Istanbul	2020-03-19	Iran
EPI_ISL_429870	hCoV-19/Turkey/HSGM-8990/2020	Turkey / Sakarya	2020-03-22	Saudi Arabia
EPI_ISL_428723	hCoV-19/Turkey/HSGM-8964/2020	Turkey / Aksaray	2020-03-22	Saudi Arabia
EPI_ISL_429871	hCoV-19/Turkey/HSGM-10241/2020	Turkey / Ankara	2020-03-23	Saudi Arabia
EPI_ISL_437317	hCoV-19/Turkey/HSGM-1027/2020	Turkey / Ankara	2020-03-27	Saudi Arabia
EPI_ISL_437329	hCoV-19/Turkey/HSGM-6204/2020	Turkey / Ankara	2020-03-19	Saudi Arabia
EPI_ISL_437331	hCoV-19/Turkey/HSGM-1014/2020	Turkey / Ankara	2020-03-25	Saudi Arabia

EPI_ISL_468160) from Pakistan. Previously, the T265I mutation was detected exclusively in the American population at a frequency of 43% making it a signature SNP for the USA, whereas it was found at very low frequency in Asia (4.8%) [15]. Likewise, a nucleotide change (C > T) was observed at position 2113 in seven Turkish samples and two isolates from Saudi Arabia, but not in any other sequence in this report.

In SARS-CoV, nsp3 has been proposed to work with nsp4 and nsp6 to induce double-membrane vesicles (DMVs), which serve as an important component of the replication/transcription complex [16]. In this report, some novel variants in the nsp3 region were also detected at positions 3903 (C > T), 5736 (C > T), and 7765 (C > A). Two of these were missense mutations (C3903T and C5736T) resulting in a changes from proline to leucine (P1213L) and alanine to valine (A1824V), respectively, and each was exclusively seen in isolates from Turkey, while C7765A was a silent mutation and was observed in 11% (9/80) of Turkish samples and only one viral isolate from Saudi Arabia (EPI_ISL_437463) (Table S2). In mutant A1824V, there is a loss of a turn and the addition of a sheet structure at position 1825 (Fig. 1h) that might have significant functional implications, whereas no substantial change in secondary structure was observed for mutant P1213L (Fig. 1i). It is important to evaluate these mutations in nsp3, as this gene has been reported to harbour many mutations that resulted

in the evolution of betacoronaviruses with extensive selection pressure [17].

Apart from frequent mutations, two mutations in the nsp4 gene (C8782T and G8653T) and one in the nsp12/RdRP gene (13,730C > T) were found to be completely unique to Turkey. A silent mutation at position 8782 (C > T) within nsp4 was found in two infected individuals (EPI_ISL_428718 and EPI_ISL_437317) from Turkey and in seven samples from Spain, where it was an infrequent mutation; however, it was present at high frequency in isolates from Oceania and North America in previous reports [8]. The G8653T mutation in the nsp4 gene, resulting in a change of methionine (M) to isoleucine (I) at position 2796 was found in 11 Turkish strains but not observed in any other viral isolates from Europe so far. However, this mutation was detected by Joshi and Paul [18] in nine Indian samples and in two isolates from Kuwait. In the nsp12/RdRP region, an alanine/valine substitution was observed at site 4489 in a single sequence (EPI_ISL_455719, Turkey/Mardin) on April 9, 2020, that was not found in any other sequence in this report. In a previous report by Maitra et al. [19], this unique mutation was also found in two infected individuals in India.

Further, our secondary structure prediction analysis showed that the M2796I mutation causes changes in the secondary structure of nsp4 in which the α -helix is replaced with a β -sheet structure at position 2795 and a turn at position 2798 (Fig. 1j), which might affect the interaction between nsp3 and nsp4 and thus the replication of the virus.

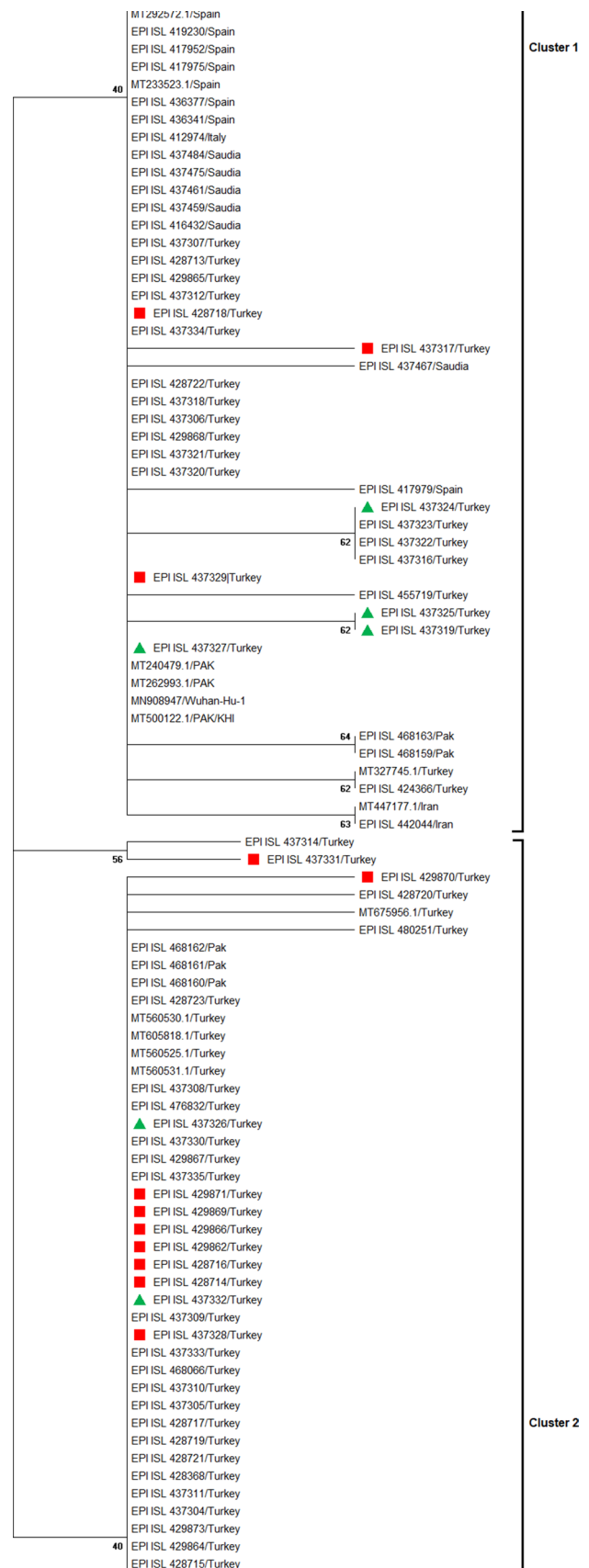
Fig. 2 Sub-tree showing the informative branch containing imported cases to Turkey from Saudi Arabia (indicated by red squares) and Iran (indicated by green triangles)

The side chain of valine is larger than that of alanine, and substitution of valine in nsp12 resulted in the potential loss of an α -helix at positions 4486, 4487, and 4488 with the addition of a β -sheet from positions 4486 to 4490 (Fig. 1k). Therefore, this substitution might have functional consequences that can potentially affect the replication and mutagenic capability of SARS-CoV-2.

Importantly, it would be interesting to investigate the effects of these substitutions in the non-structural proteins, as one of a previous study has suggested that an alanine-to-valine substitution in the non-structural protein NS2A of Zika virus impairs viral RNA synthesis and results in attenuation of the virus [20]. Similarly, a valine substitution in the RdRP protein in Indian SARS-CoV-2 isolates has been shown to cause a structural alteration that impairs packing of the protein [21]. Thus, functional characterization of the mutations investigated in our study needs to be carried out to understand their role and to develop strategies for vaccine design.

We found that 13 viral isolates from Turkey harboring 23,403 A > G, 3037 C > T, and 14,408 C > T mutations were from individuals who had a history of travel from Saudi Arabia (Table 2). Of these, six cases were reported from the city of Ankara while the remaining ones were from Aksaray, Sakarya, Tekirdag, Kocaeli, Kastamonu, Konya and Afyon. Similarly, six samples were from individuals who had a history of travel from Iran and most of these cases (5/6) were from Istanbul and were listed in GISAID. Among Iran-travel-linked isolates, four samples (EPI_ISL_437319, EPI_ISL_437324, EPI_ISL_437325, and EPI_ISL_437327) contain the mutations (G1397A, T28688C and G29742T, a Europe-based introduction), while the remaining two samples (EPI_ISL_437326 and EPI_ISL_437332) lacked these mutations (Table 2). However, the G1397A and T28688C substitutions were also observed in one patient who had travelled to Taiwan. This suggests that COVID-19 infection spread to Turkey from multiple countries, especially from Saudi Arabia and Iran. In a similar analysis, Eden et al. [22] provided evidence of SARS-CoV-2 being exported to Australia from Iran. Our findings may also contribute to a better understanding of the diversity of circulating SARS-COV-2 strains and origin of imported cases in Turkey.

Phylogenetic analysis of SARS-COV-2 genome sequences suggested that the Turkish strain is closely related to isolates from Saudi Arabia, suggesting a common origin (Fig. 2). However, the sequences from Turkey were dispersed throughout the phylogenetic tree, indicating multiple independent introductions into the country. In the phylogram, two distinct clades were categorized as cluster 1



and cluster 2. Cluster 1 contained the mutations 1397G > A, 11,083G > T, 28,688 T > C, and 29,742G > T, while the second cluster had 23,403 A > G, 3037 C > T, and 14,408 C > T as the dominant mutations. Moreover, the majority of viral isolates in Turkey showed L type characteristics and formed a monophyletic clade, while the S type was present in limited numbers. An SNP with “T” at position 28,144 encoding leucine was classified as L type, while “C” at this position encoding serine is referred as S type. It has been suggested that the L type is more aggressive and contagious than the S type [12]. Thus, it appears that the L type is predominant in the Turkish population. In addition, samples having travel connection to Saudi Arabia and Iran showed a monophyletic origin within their respective clusters in the phylogenetic tree.

In conclusion, functional characterization of novel mutations investigated in our study needs to be carried out to understand the exact role of these variations. Furthermore, awareness of the above-mentioned mutations might be useful for the identification of less-virulent strains and the development of vaccines against a large repertoire of strains. Phylogenetic and transmission analysis revealed that the spread of SARS-CoV-2 to Turkey was due to multiple independent introductions and that viral isolates from Saudi Arabia and Turkey are closely related to one another.

Compliance with ethical standards

Competing interests The authors declare that they have no competing interests.

References

- Gorbalenya A et al (2020) The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiol* 5:536–544
- WHO. Transmission of SARS-CoV-2: implications for infection prevention precautions (2020). <https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2> implications-for-infection-prevention-precautions. Accessed 5th May, 2020
- World Health Organization. Coronavirus disease 2019 (COVID-19) Situation Report-95. Available From: URL: https://www.who.int/docs/default-source/coronaviruse/situationreports/20200424-sitrep-95-covid-19.pdf?sfvrsn=e8065831_4.
- Ministry of Health of Turkey. Current status in Turkey. 2020. <https://covid19.saglik.gov.tr> (Accessed May 20, 2020)
- Kumar S, Stecher G, Tamura K, MEGA7 (2016) Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7):1870–4
- Ashok Kumar T (2013) CFSSP: Chou and Fasman secondary structure prediction server. *Wide Spectrum* 1:15–19. <https://doi.org/10.5281/zenodo.50733>
- Banerjee S, Seal S, Dey R, Bhattacharjee P (2020) Mutational spectra of SARS-CoV-2 orf1ab polyprotein and Signature mutations in the United States of America. *bioRxiv*
- Pachetti M, Marini B, Benedetti F, Giudici F, Mauro E et al (2020) Emerging SARS-CoV 2 mutation hot spots include a novel RNA - dependent - RNA polymerase variant. *J Trans Med* 18(179):1–9. <https://doi.org/10.1186/s12967-020-02344-6>
- Ayub MI (2020) Reporting two SARS-CoV-2 strains based on a unique trinucleotide-Bloc mutation and 351 their potential pathogenic difference. Preprints. <https://doi.org/10.20944/preprints202004.0337.v1>
- Mercatelli D, Giorgi FM (2020) Geographic and genomic distribution of SARS-CoV-2 mutations. Preprints 338(2020):2020040529. <https://doi.org/10.20944/preprints202004.0529.v1>
- Benvenuto D, Angeletti S, Giovanetti M, et al (2020) Evolutionary analysis of SARS-CoV-2: how mutation of Non-Structural Protein 6 (NSP6) could affect viral autophagy [published online ahead of print]. *J Infect* S0163–4453(20)30186–9. <https://doi.org/10.1016/j.jinf.2020.03.058>.
- Samyuktha V, Kumar V N (2020) Emergence of RBD and D614G Mutations in Spike Protein: An Insight from Indian SARSCoV-2 Genome Analysis. <https://doi.org/10.20944/preprints202006.0032.v1>
- Ueda MT, Kurosaki Y, Izumi T, Nakano Y, Oloniniyi OK et al (2017) Functional mutations in spike glycoprotein of Zaire ebolavirus associated with an increase in infection efficiency. *Genes Cells* 22:148–159
- Kurosaki Y, Ueda MT, Nakano Y, Yasuda J, Koyanagi Y, Sato K, Nakagawa S (2018) Different effects of two mutations on the infectivity of Ebola virus glycoprotein in nine mammalian species. *J Gen Virol* 99:181–186
- Rouchka EC, Chariker JH, Chung D (2020) Phylogenetic and Variant Analysis of 1,040 SARS-CoV-2 Genomes. Preprints. <https://doi.org/10.20944/preprints202005.0396.v1>
- Sakai Y, Kawachi K, Terada Y, Omori H, Matsuura Y, Kamitani W (2017) Two-amino acids change in the nsp4 of SARS coronavirus abolishes viral replication. *Virology* 510:165–174
- Wu C, Liu Y, Yang Y, et al (2020) Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B*
- Joshi A, Paul S (2020) Phylogenetic analysis of the novel coronavirus reveals important variants in Indian strains. Preprint <https://doi.org/10.1101/2020.04.14.041301>
- Maitra A, Sarkar MC, Raheja H et al (2020) Mutations in SARS-CoV-2 viral RNA identified in Eastern India: Possible implications for the ongoing outbreak in India and impact on viral structure and host susceptibility. *J Biosci*. <https://doi.org/10.1007/s12038-020-00046-1>
- Márquez-Jurado S, Nogales A, Ávila-Pérez G, Iborra FJ, Martínez-Sobrido L, Almazán F (2018) An alanine-to-valine substitution in the residue 175 of Zika virus NS2A protein affects viral RNA synthesis and attenuates the virus in vivo. *Viruses* 10(10):547
- Chand GB, Banerjee A, Azad GK (2020) Identification of novel mutations in RNA-dependent RNA polymerases of SARS-CoV-2 and their implications on its protein structure. *bioRxiv*
- Eden JS, Rockett R, Carter I, Rahman H, De Ligt J et al (2020) An emergent clade of SARS-CoV-2 linked to returned travellers from Iran. *Virus Evol* 6(1):1–4. <https://doi.org/10.1093/ve/veaa027>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.