AIJR Preprints
Section: Coronavirus
Article Id: 251, Version: 1, 2020
URL: https://preprints.aijr.org/index.php/ap/preprint/view/251
{Click on above link to see the latest available version of this article}

NOT PEER-REVIEWED

Version 1: Received: 14 October 2020 / Approved: 15 October 2020 / Online: 15 October 2020

Genetic variation of SARS-CoV-2 circulating worldwide and its association for altering disease fatality

Rimjhim Dasgupta

4NBIO, 2502, Glen Classic, Hiranandani Gardens, Powai, Mumbai 400076

Author's email: **********

ABSTRACT

The emergence of SARS-CoV-2 has resulted in > 36,361,054 infections and > 1,056,186 deaths worldwide. Using publicly available genome sequences of patient samples from different geographical regions, a study has been conducted to co-relate mutational frequency with disease transmission and fatality rate. Seven hundred genome sequences were randomly chosen from different countries. The regions of the genome encoding structural proteins Spike (S), Nucleocapsid (N), envelop (E) and Membrane (M) proteins and ORF8 were studied here. Through Insilco approach, this study showed that several evolutionary conserved amino acid residues underwent mutations. Some of these mutations are common in multiple geographies. Quite a few region-specific mutations are also identified. This study highlights that mutational rate is proportional to disease transmission and inversely proportional to disease fatality. The changes in the conserved residues have significant implication on the stability of the proteins and subsequent interaction, which are essential for virus propagation. This provides a better understanding of the genetic variation in SARS-CoV-2 across the countries and its association with reducing disease fatality.

Keywords: SARS-CoV2, disease fatality, mutation

1 Introduction

The COVID-19 pandemic caused by a novel 2019 SARS coronavirus, known as SARS-CoV-2, is rapidly spreading worldwide after one month of the initially identified case on December 2019, in Wuhan city, China [1]. The genome sequence study has revealed that SARSCoV-2 is a member of the genus Beta-coronavirus and belongs to the subgenus Sarbecovirus that includes SARS-CoV while MERS-CoV belongs to a separate subgenus, Merbecovirus [2, 3]. Epidemiological data suggests that SARS-CoV-2 had spread widely from the city of Wuhan in China [4] after its zoonotic transmission originating from bats via the Malayan pangolins [5]. It has spread over 200 countries and infected millions of people worldwide. As the number of the positive cases increasing drastically, the World Health Organization (WHO) raised the importance of understanding genetic changes through mutation that could have occurred in the SARS-CoV-2. The SARS-CoV-2 genome is composed of approximately 30,000 nucleotides [5]. The genome includes a variable number (from 6 to 11) of open reading frames (ORFs) [6]. The first ORF (ORF1ab) representing approximately 67% of the entire genome encodes 16 non-structural proteins (nsps), while the remaining ORFs encode accessory proteins including ORF8 and structural proteins. The four major structural proteins are the spike surface glycoprotein (S), small envelope protein (E), membrane protein (M), and nucleocapsid (N) protein [6].
Searching for mutations and their evolutionary conservation while the virus continues to spread, can offer opportunities for a better understanding of virus evolution, biopathology, and transmission.

How to Cite:

Rimjhim Dasgupta, "Genetic variation of SARS-CoV-2 circulating worldwide and its association for altering disease fatality". *AIJR Preprints*, 251, Version 1, 2020.

Having this motivation, considering Wuhan based genome NC_045512.2 as reference, this study attempted to understand worldwide common hotspots, varying mutational frequency from region to region and its association for reducing disease fatality.

2 Methods

In the present report we randomly chosen ~ 700 genome sequences (sourced from Global Initiative on Sharing All Influenza, GISAID, https://www.gisaid.org/ and NCBI) worldwide. The study is based on sequence alignments and used NCBI BLAST, and CLUSTAL OMEGA. Wuhan isolate, SARS-CoV-2 sequence NC_045512.2 (length 29903 nt) was used as a reference sequence and for sequence comparisons.

Mutational rate has been calculated by counting total number of missense mutations in all structural proteins (Spike, Nucleocapsid, Envelope and Membrane proteins) and ORF8 (considered in this study) in each genome followed by dividing the total number of genome sequences undergone mutation. The regions in which sufficient genome sequences (minimum 10 and maximum 60 genome sequences) are not available during this study has been excluded.

Fatality rate of each country was measured using cumulative death (https://covid19.who.int/) divided by total number of affected cases and then multiplied by hundred.

3 Results

All the different types of missense mutations in S, N, E and ORF8 proteins have been recorded and placed in tables (1, 2, 3, 4). S protein contains maximum number of mutational hotspots (Table 1). All the countries studied here, have common D614G mutation for S protein. India has maximum number of different types of substitution in S protein. Similarly, N protein has few mutations which are common in many countries (table 2). ORF8 (table 3) and E protein (table 4) have very few mutations whereas M protein doesn't have any mutation for the genome sequences studied here.

L5F substitution in S protein was found in some genome sequences from the patient samples of Bangladesh, India, Italy (very few), Philippines and USA. This (L5F mutation) is mostly coupled with D614G substitution. In addition, L54F substitution was found in few sequences of patient samples from India which is coupled with D614G mutation. The receptor-binding domain or RBD of the S protein of SARS-CoV-2 lies between amino acids 330 and 583. Maximum number of mutations were found in N terminal domain followed by RBD and C terminal domain (Table 1).

S194L, S202N, R203G, G204R mutations are abundantly found in N (nucleocapsid) protein. These mutations are found in sequences from multiple countries. Again, like S protein, India has maximum number of different types of substitution in N protein (table 2). T205I mutation is only noted in all N protein sequences from Iraq. P67S or S235F are found in many sequences (N protein) from USA and these are not coupled with any of the other mutations such as S194L, S202N, R203G and G204R.

For ORF8, very few mutational hotspots were found in the sequences studied here. USA has maximum number of substitutions in ORF8 (table 3). L84S mutation is common among the sequences of patient samples from India and USA.

This study reports very few mutations in E protein (table 4). It did not find any common mutational hotspot in sequences from multiple countries.

Next, mutation rate was calculated and plotted with respect to different countries (figure 1). It shows that Bangladesh, India and Philippines have the highest mutational frequency whereas Australia, Italy, Egypt have least.

Then fatality rate was calculated using WHO cumulative death (https://covid19.who.int/) and plotted against mutational rate. This study reports that they are inversely related (figure 2).

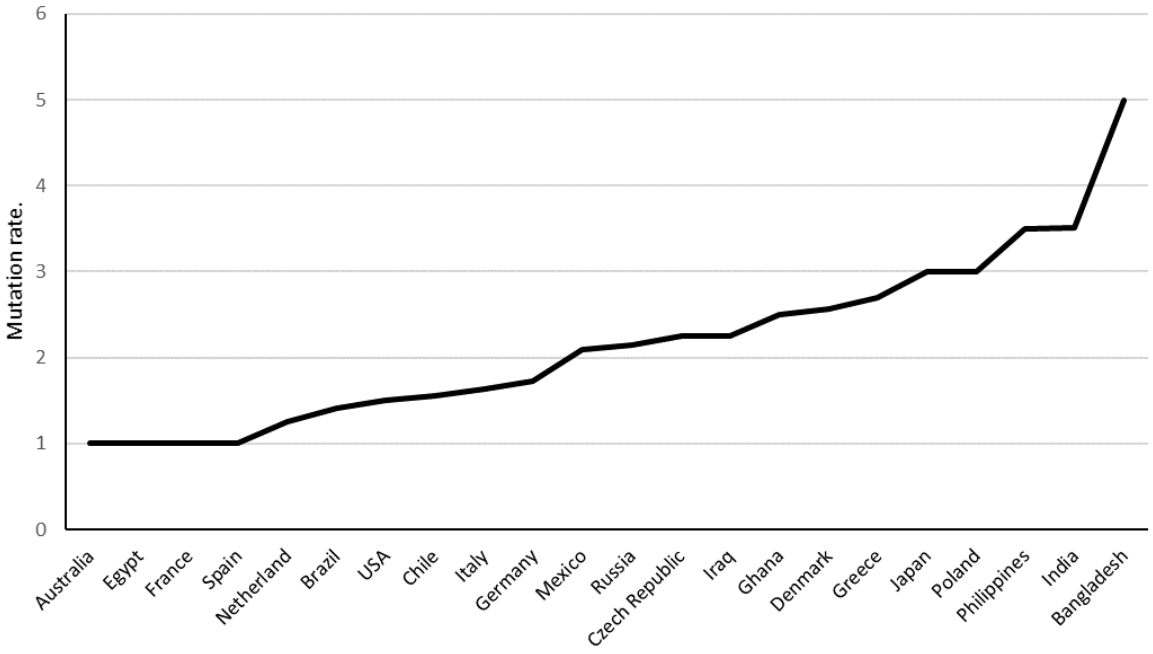

Figure 1: Mutational rate has been plotted with respect to country. It has been calculated by counting number of missense mutations in all structural proteins and ORF8 in each followed by dividing the total number of genome sequences undergone mutation. We excluded regions in which sufficient genome sequences (minimum 10 and maximum 60 genome sequences) are not available during this study

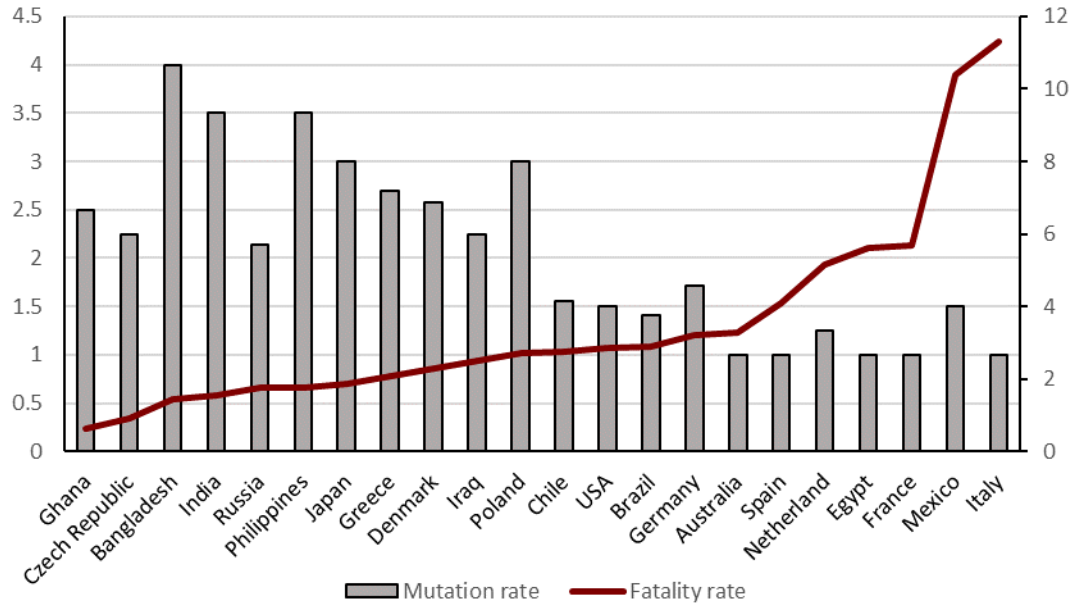

Figure 2: Mutational and fatality rates have been plotted. Fatality rate has been calculated using cumulative death (https://covid19.who.int/) divided by total number of affected cases and then multiplied by hundred

4 Discussion

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is being intensively studied, particularly its evolution, in the increasingly available sequences with classical phylogenetic tree representation. This study reports certain amino acid variations in structural proteins and ORF8 and its possible implication for altering structure, function, infectivity and fatality.

Table 1: The different types of missense mutations in S (Spike) protein for different countries (vertically) are placed in tabular form. Single letter amino acid code has been used and placed horizontally in table; + is used for presence and - is used for absence of respective mutation

Mutation →	L5F	S11F	R21T	Y28H	I39V	H49Y	L54F	T76I	R78M	Y145Stop	H146Y/R	W152L	S162I	L176F	N189D	I197V	G261D	A288T	Q314R	R407I	Y452F	A522V	T572I	E583D	D614G	V622F	Q677H	S884F	A930V	T940A	V1122L	G1124V	R1185H	M1237I
Country ↓																																		
Australia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Bangladesh	+	-	-	-	+	+	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-
Brazil	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Chile	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Czech Republic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Denmark	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-
Egypt	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
France	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Germany	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Greece	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-
India	+	-	-	+	-	-	+	-	+	+	-	+	+	-	-	-	-	-	-	+	-	-	+	+	+	-	+	-	+	-	-	+	-	-
Iran	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Iraq	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-
Italy	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Jamaica	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Japan	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Jordan	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Mexico	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Morocco	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Netherland	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-
Philippines	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Poland	-	-	-	-	-	-	-	-	-	-	-	-	-	+		-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Puerto Rico	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Russia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Spain	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Sri Lanka	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Taiwan	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-
Tunesia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
UK	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
USA	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	+

Coronavirus entry into host cells is mediated by the transmembrane spike (S) protein that forms homotrimers protruding from the viral surface [7]. S protein comprises two functional subunits, responsible for binding to the host cell receptor (S1 subunit) and fusion of the viral and cellular membranes (S2 subunit). For many Coronaviruses, S is cleaved at the boundary between the S1 and S2 subunits, known as furin cleavage site and remain non-covalently bound in the prefusion conformation. This region is reported to be the most potent and indispensable for viral attachment and entry into host system [8]. This study reports that amino acid at position 614 (a residue close to furin cleavage site) underwent mutation (D614G) which is common across all geographies. Earlier study reported that amino acid at position 614 occurs at an internal protein interface of the viral spike, and the presence of G at this position destabilises a specific conformation [9]. There are several other miss-sense mutations in S protein (table 1). All these mutations could be classified as stabilizing and destabilizing based on the free-energy changes. L5F mutation in sequences from Bangladesh, India, Italy (very few), Philippines and USA patient samples and L54F mutation in sequences from India, are mostly coupled with D614G substitution. Both of the amino acids (leucine and phenylalanine) are non-polar, but phenylalanine has a benzoic ring in the side chain which may stiffen the secondary structure by means of aromatic-aromatic, hydrophobic or stacking interactions. Earlier report in Non- Structural Protein 6 (NSP6) by amino acid change stability (ACS) analysis showed that this (leucine to phenylalanine) leads to a lower stability of the protein structure [10]. There are multiple residues in receptor binding domain (RBD) which play important role in binding to ACE2 (angiotensin converting enzyme 2) receptor [11]. This study reports multiple point mutations in neighbouring amino acids of ACE2 binding residues. All

these potentially play important role for increasing person to person transmission by altering the affinity to ACE2, stability of protein and interaction with neighbouring molecules.

Table 2: The different types of missense mutations in N (nucleocapsid) protein for different countries (vertically) are placed in tabular form. Single letter amino acid code has been used and placed horizontally in table; + is used for presence and - is used for absence of respective mutation

Mutation →	P6T	A12G	P13L	D22Y/N	S33I	G35W	P67S	T121I	A134V	L139F	R149L	S180T	S183Y	S186F	S194L	S197L	S202N	R203K	G204R	T205I	S235F	T282I	P344S	D347H	D348Y	D371Y
Country ↓																										
Bangladesh	-	-	-	-	-	-	-	-	-	-	-	+		-	-	-	-	+	+	-	-	+	-	-	-	-
Brazil	-	-	-	-	-	-		-	-	-	-	-		-	-	-	-	+	+	-	-	-	-	-	-	-
Chile	-	-	-	-	-	-		-	-	-	-	-		-	-	+	-	+	+	-	-	-	-	-	-	-
Czech Republic	-	-	-	-	-	-		-	-	-	-	-		-	-	-	-	+	+	-	-	-	-	-	-	-
Denmark	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Germany	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Greece	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
India	+	-	+	+	+	-	-	-	+	+		-	-		+	-	+	+	+	-	-	-	+	-	+	-
Iran	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+		-	-	-	-	-	-	-	-	-	-
Iraq	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-
Israel	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Italy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Jamaica	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Japan	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Mexico	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
Morocco	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-
Philippines	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Poland	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Puerto Rico	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Russia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Russia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Russia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Russia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Sri Lanka	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Taiwan	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Tunesia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
UK	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
USA	-	-	-	-	-	+	+	-	-	-	-	-	+	-	-	-	-	+	+	-	+	-	-	-	-	+

The nucleocapsid (N) protein is an important structural protein for the coronaviruses. It is highly abundant in the viruses. Its function involves, entering into the host cell, binding to the RNA, and forming the ribonucleoprotein core. It consists of RNA binding domain (RBD; residues 44-180) in the N-terminal region (N) of the protein, linker peptide (residues 181-246), the dimerization domain (DD; residues 247-364) in the C-terminal region [12]. Three disordered regions were reported on 1) N terminal (residues 1-43), 2) linker peptide, and 3) C terminal end (residues 365-419) [12]. This study reports multiple point mutations in linker region (S194L, S202N, R203K and G204R) and these are found to be common for many countries (Table 2). Earlier report pointed out that these positions are evolutionary conserved [13]. It was also reported that multiple disordered regions facilitate the N protein to transiently bind to different partners and maintain a correct conformation [12]. The hotspots reported here, surround with GSK3 phosphorylation 'SRGTS' (amino acid position 202-206) and CDK phosphorylation 'SPAR' (amino acid position 206-209) motifs [14]. Most possibly the substitutions alter the binding affinity to attain the stable conformation and simultaneously contribute towards enhancing transmission rate.

This study reports very few substitutions in ORF8 (table 3). L84S was noted in sequences from USA and India. I39V was found in ORF8 sequences from Bangladesh. Earlier study showed that ORF8 along with N and ORF3b are potent interferon antagonist, in the early stages of SARS-CoV-2 infection [15].

It hinders the host's antiviral response and then benefit virus replication by delaying the release of IFNs [15]. For L84S, Leu is non-polar hydrophobic whereas Ser is polar amino acid. Ser residue undergoes phosphorylation and possibly play an important role for reducing anti-interferon activity and hence the disease severity.

Table 3: The different types of missense mutations in ORF8 protein for different countries are placed in tabular form. Single letter amino acid code has been used; + is used for presence and - is used for absence of respective mutation

Mutation →	F3I	A14S	S24L	I39V	V62L	A65V	L84S
Country ↓							
Bangladesh	-	-	-	+	-	-	-
India	-	+	-	-	-	-	+
USA	+	-	+	-	+	+	+

The envelope (E) protein is a small, integral membrane protein involved in several aspects of the virus' life cycle, such as assembly, budding, envelope formation, and pathogenesis. The SARS-CoV E protein consists of three domains, i.e. the amino (N)-terminal domain (residues ~1-8), the transmembrane domain (TMD, residues 9 -38)), and the carboxy (C)-terminal domain (residues 39-75) [16]. We found very few genetic variation (L21F, I33T, L49M, V62F) in E protein and none of these substitutions are common worldwide (table 4). L21F potentially interfere the oligomerisation of SARS-CoV2 as it was shown earlier that V25F hampers the oligomerisation of SARS-CoV E [16]. V62F mutation in C-terminal domain can potentially interfere the interaction with target proteins, thereby altering the host-cell processes required for viral infection [12].
This analysis does not find any substitution in Membrane (M) protein in the genome sequences of ~700 patient samples of different geographical regions. This implies that M protein is comparably stable and supports our earlier study [13].

Table 4: The different types of missense mutations in E (envelope) protein for different countries are placed in tabular form. Single letter amino acid code has been used; + is used for presence and - is used for absence of respective mutation

Mutation →	V5A	L21F	I33T	L49M	V62F
Country ↓	-	-	-	-	-
Brazil	+	-	-	-	-
Philippines	-	+	-	-	-
Greece	-	-	-	+	-
India	-	-	-	-	+
Iran	-	-	+	-	-

We noticed D614G for S and S194L, S202N, R203K, G204R for N proteins are common mutations for all most all countries except China (at least the genome sequences studied here). While looking into the total affected cases and fatality rate from WHO data (https://covid19.who.int/), China shows comparatively less transmission rate but more fatality.

To understand the possible implication of these mutations, first we plotted mutational frequency (rate) with respect to different countries (figure 1). It shows that Bangladesh, India and Philippines have the

highest mutational frequency. Then we tried to relate mutational frequency with fatality rate. Very interestingly, the result shows that Bangladesh, India and Philippines have less fatality whereas Italy, Mexico, France have less mutation but more fatality. In other words, the fatality rate is decreasing with increasing mutation rate or vice versa (figure 2). This finding is consistent with earlier studies on NSP6, S protein and RNA-dependent RNA polymerase (RdRP) [17, 18].

5 Conclusion

This study reports several missense mutations for around seven hundred SARS-CoV2 genome sequences studied here of patients from diverse geographical-locations. Within a very short time frame, the virus evolved rapidly. Quite a few common mutations across geographies in S protein, N protein, and also in ORF8 indicate that these variations were mostly carried by the patients with travel history and then transmitted to neighbouring or family members. The potential implication of these may link to enhancing the infectivity, easy person to person transmission. There are many variations those were evolved within the country as we found quite a few country specific mutations. Although the clinical significance of the observed mutations is not yet available, our findings lay the groundwork to understand the impact of SARS-CoV2 mutations on disease severity. This study also warrants the importance of sequencing the whole genome of SARS-CoV-2 after several passages and key mutations should be used for the effective drug designing and treatment options. All together our findings make us optimistic that the disease severity will diminish as we move along with time and more genomic variations.

6 Competing Interests

The author declares no competing financial interests.

References

[1] Farkas C, Fuentes-Villalobos F, Garrido JL, Haigh J, Barría MI. 2020. Insights on early mutational events in SARSCoV-2 virus reveal founder effects across geographical regions. PeerJ 8:e9255. https://doi.org/10.7717/peerj.9255

[2] Lu R, Zhao X, Li J, Niu P, Yang B. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 395, pp. 565–574, 2020

[3] Zhu N, Zhang D, Wang W, Li X, Yang B et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New Eng. J. Med*. 382, pp. 727–733, 2020

[4] Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 368, pp. 395–400, 2020

[5] Zhang T, Wu Q and Zhang Z. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr. Biol.* 30, pp. 1346–1351, 2020

[6] Song, Z.; Xu, Y.; Bao, L.; Zhang, L.; Yu, P.; Qu, Y.; Zhu, H.; Zhao, W.; Han, Y.; Qin, C. From SARS to MERS, Thrusting Coronaviruses into the Spotlight. *Viruses* 2019, *11*, 59

[7] Tortorici, M.A., and Veesler, D. Structural insights into coronavirus entry. *Adv. Virus Res*. 105, pp. 93–116. 2019.

[8] Alexandra C. Walls, Young-Jun Park, M. Alejandra Tortorici, Abigail Wall, Andrew T. McGuire, David Veesler. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein; *Cell* 180, pp. 281–292, 2020. https://doi.org/10.1016/j.cell.2020.02.05827

[9] Becerra-Flores M, Cardozo T. SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate. *Int J Clin Pract*. 00: e13525, 2020. https://doi.org/10.1111/ijcp.13525

[10] Domenico Benvenuto, Ayse Banu Demir, Marta Giovanetti, Martina Bianchi, Silvia Angeletti, Stefano Pascarella, Roberto Cauda, Massimo Ciccozzi, Antonio Cassone. Evolutionary analysis of SARS-CoV-2: how mutation of Non-Structural Protein 6 (NSP6) could affect viral autophagy, *Journal of Infection* 81, e24–e27, 2020. https://doi.org/10.1016/j.jinf.2020.03.058

[11] Dasgupta, R. Comparative genomics of receptor binding domains of S protein and host receptor interaction in COVID-19 patient. International Journal of Creative Research Thoughts (IJCRT), Volume.8, Issue 8, pp.2571-2575, August 2020.

[12] Weihong Zeng, Guangfeng Liu, Huan Ma, Dan Zhao, Yunru Yang, Muziying Liu, Ahmed Mohammed, Changcheng Zhao, Yun Yang, Jiajia Xie, Chengchao Ding, Xiaoling Ma, Jianping Weng, Yong Gao, Hongliang He, Tengchuan Jin, Biochemical characterization of SARS-CoV-2 nucleocapsid protein. *Biochemical and Biophysical Research Communications* ,527, pp. 618-623, 2020

[13] Dasgupta R. Mutational Analysis Of Structural Proteins In Sars-Cov-2 Viral RNA Of Covid-19 In India. Research Journal Of Life Sciences, Bioinformatics, Pharmaceutical and chemical Sciences. Vol 6, Issue 5, 2020. DOI: 10.26479/2020.0605.01

[14] Surjit M, Kumar R, Mishra RN, Reddy MK, Chow VTK. The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated localizes in the cytoplasm by 14-3-3-mediated translocation. *J. Virol*. 79, pp. 11476–11486, 2005

[15] Li JY, Liao CH, Wang Q, Tan YJ, Luo R, Qiu Y, Ge XY. The ORF6, ORF8 and nucleocapsid proteins of SARS-CoV-2 inhibit type I interferon signaling pathway. Virus Res. 2020 Sep;286:198074. https://doi.org/10.1016/j.virusres.2020.198074

[16] Dewald Schoeman and Burtram C. Fielding. Coronavirus envelope protein: current knowledge. *Virology Journal* 16,69, 2019. https://doi.org/10.1186/s12985-019-1182-0

[17] Jalali S, Bhadra B, Dasgupta S. Significance of Mutation Rate of Structural and Non-Structural Proteins of SARS-Cov-2 Showed with Lower Death Rate of COVID-19. Sci J Biol. 2020;3(1): 017-022. https:// dx.doi.org/10.37871/sjb.id18

[18] Samir Kumar Patra. "Perspective on Accelerating the Mutation Rate of SARS-CoV-2 for a Better Way of COVID-19 Treatment; Enhanced Mutation Therapy of COVID-19". AIJR Preprints, 62, version 1, 2020. https://preprints.aijr.org/index.php/ap/preprint/view/62