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.

International Immunopharmacology 105 (2022) 108565

Contents lists available at ScienceDirect

International Immunopharmacology

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

A review on evolution of emerging SARS-CoV-2 variants based on spike glycoprotein

Nimisha Ghosh [a,b,1], Suman Nandi [c,1], Indrajit Saha [c,1,*]

[a] Faculty of Mathematics, Informatics, and Mechanics, University of Warsaw, Warsaw, Poland
[b] Department of Computer Science and Information Technology, Institute of Technical Education and Research, Siksha `O' Anusandhan (Deemed to be) University, Bhubaneswar, Odisha, India
[c] Department of Computer Science and Engineering, National Institute of Technical Teachers' Training and Research, Kolkata, West Bengal, India

ARTICLE INFO

Keywords:
COVID-19
Mutations
SARS-CoV-2 genomes
Spike glycoprotein
Virus strains

ABSTRACT

Since the inception of SARS-CoV-2 in December 2019, many variants have emerged over time. Some of these variants have resulted in transmissibility changes of the virus and may also have impact on diagnosis, therapeutics and even vaccines, thereby raising particular concerns in the scientific community. The variants which have mutations in Spike glycoprotein are the primary focus as it is the main target for neutralising antibodies. SARS-CoV-2 is known to infect human through Spike glycoprotein and uses receptor-binding domain (RBD) to bind to the ACE2 receptor in human. Thus, it is of utmost importance to study these variants and their corresponding mutations. Such 12 different important variants identified so far are B.1.1.7 (Alpha), B.1.351 (Beta), B.1.525 (Eta), B.1.427/B.1.429 (Epsilon), B.1.526 (Iota), B.1.617.1 (Kappa), B.1.617.2 (Delta), C.37 (Lambda), P.1 (Gamma), P.2 (Zeta), P.3 (Theta) and the recently discovered B.1.1.529 (Omicron). These variants have 84 unique mutations in Spike glycoprotein. To analyse such mutations, multiple sequence alignment of 77681 SARS-CoV-2 genomes of 98 countries over the period from January 2020 to July 2021 is performed followed by phylogenetic analysis. Also, characteristics of new emerging variants are elaborately discussed. The individual evolution of these mutation points and the respective variants are visualised and their characteristics are also reported. Moreover, to judge the characteristics of the non-synonymous mutation points (substitutions), their biological functions are evaluated by PolyPhen-2 while protein structural stability is evaluated using I-Mutant 2.0.

1. Introduction

The ongoing wave of COVID-19 caused by SARS-CoV-2 virus was first identified in the city of Wuhan, China during December 2019. Since then, the virus has spread very rapidly and has affected millions of people worldwide. SARS-CoV-2 is a positive stranded RNA virus with a length of about 30 kb encompassing non-structural and structural proteins. Spike glycoprotein, a structural protein present on the virus surface plays an important role in binding with ACE2. This RNA virus can make a replica of its own after binding with the host cell, thereby causing several mutations [24]. Whenever the mutation is significant, the structure of the virus changes, resulting in a new variant or lineage[2] of the virus [38]. Motivated by this observation, in this study we have performed a competitive analysis of several variants of SARS-CoV-2. The mutation of SARS-CoV-2 is happening over time, thereby resulting in new variants. Whenever a new variant emerges, it can be called as an "emerging variant" which have some potential consequences viz. increase in transmissibility, morbidity as well as mortality. It is to be noted that the different variants have some unique as well as some common mutations. In this regard, there are 12 important variants as declared by W.H.O[3] and 84 unique mutations that are reported in this work. Some of these variants have been categorised as either variants of concern, variants of interest or variants under monitoring based on their transmissibility, immunity and infection severity[4]. As of now, the variants of

* Corresponding author.
E-mail address: indrajit@nitttrkol.ac.in (I. Saha).
[1] Equally contributed.
[2] https://cov-lineages.org/lineages/lineage_B.1.html.
[3] https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/.
[4] https://www.ecdc.europa.eu/en/covid-19/variants-concern.

https://doi.org/10.1016/j.intimp.2022.108565
Received 8 November 2021; Received in revised form 19 January 2022; Accepted 20 January 2022
Available online 29 January 2022

concern are Alpha (B.1.1.7), Beta (B.1.351), Gamma (P1), Delta (B.1.617.2) and Omicron (B.1.1.529)3. The two common features which mark any variant of concern are multiple mutations in Spike glycoprotein as compared to B.1 which is also known as the "wild-type" (with D614G and no other Spike glycoprotein changes) as well as at least one mutation in receptor binding domain (RBD) of Spike glycoprotein4. Apart from the variants of concern, the variant of interest is Lambda (C.37) while the variants under monitoring are Eta (B.1.525), Iota (B.1.526) and Kappa (B.1.617.1). Other variants include Epsilon (B.1.427/B.1.429), Zeta (P.2) and Theta (P.3). The intention of our study is to help the researchers in understanding the significance of such variants. .

Spike glycoprotein with a length of 1273 aa covers the SARS-CoV-2 surface. This protein consists of two functional subunits: S1 which is responsible for receptor binding and S2 which is responsible for membrane fusion [42]. The N-terminal domain and the receptor binding domain (RBD) are the major two domains of S1 subunit while fusion peptide (FP), heptapeptide repeat sequence 1 (HR1), heptapeptide repeat sequence 2 (HR2), transmembrane (TM) domain and cytoplasm domain are covered by S2 subunit. In the S1 subunit, RDB is responsible for binding with angiotensin-converting enzyme 2 (ACE2) cell receptor [15]. After binding with cell receptor, TM protease serine 2 (TMPRSS2) on the receptor cell activates the Spike glycoprotein. Whenever the S1 subunit binds to the ACE2 host cell receptors, then the S2 subunits perform two major conformational changes to complete the virus fusion to the cell membrane.

Considering the aforementioned analysis, in this work we have performed multiple sequence alignment of 77681 SARS-CoV-2 genomes of 98 countries over the period from January 2020 to July 2021 using MAFFT [19] followed by phylogenetic analysis to analyse the mutations in Spike glycoprotein. 12 different important variants identified so far are Alpha, Beta, Eta, Epsilon, Iota, Kappa, Delta, Lambda, Gamma, Zeta, Theta and Omicron. These variants have 84 unique mutations and include some notable mutations like K417N, L452R, S477N, T478K, E484K/Q, N501Y, D614G, P681H/R, Y144-, H69- and V70-. Furthermore, the characteristics of the variants are elaborately discussed along with their specific mutations. Thereafter, the individual evolution of these mutation points are visualised along with their evolution in the respective variants. Moreover, the characteristics of the non-synonymous mutation points (substitutions) are judged by evaluating their biological functions by considering the sequences and using PolyPhen-2 while I-Mutant 2.0 evaluates the protein structural stability. Thus, this work provides a comprehensive review of the emerging variants and the characteristics of the corresponding mutation points along with the effects of vaccine and therapeutics on the variants.

2. Materials and methods

In this section, dataset collection for the SARS-CoV-2 genomes are elaborated followed by the proposed pipeline.

2.1. Data preparation

For multiple sequence alignment and phylogenetic analyis, 77681 global SARS-CoV-2 genomes are collected from Global Initiative on Sharing All Influenza Data (GISAID)[5] and the Reference Genome (NC 045512.2)[6] is collected from National Center for Biotechnology Information (NCBI). The SARS-CoV-2 sequences are mostly distributed globally from January 2020 to July 2021. Moreover, to map the protein sequences and changes in the amino acid for SARS-CoV-2, protein PDB are collected from Zhang Lab[7] and are then used for modelling to identify the structural changes. All these analysis are performed on High Performance Computing facility of NITTTR, Kolkata while MATLAB R2021a is used for checking the amino acid changes.

2.2. Pipeline of the work

This study is carried out according to the pipeline as given in Fig. 1 (a). Initially, 77681 global SARS-CoV-2 genomes are considered for multiple sequence alignment using MAFFT followed by their phylogenetic analysis using Nextstrain. Once the aforementioned analysis is over, the different known mutations in the Spike glycoprotein pertaining to the important SARS-CoV-2 variants are identified as shown in Fig. 1 (b) while the different domains are shown in Fig. 1(c). The entropy of the genomic coordinates of these mutation points are also calculated to show the evolution of the different variants. The entropy is calculated as follows:

$$\lambda = ln5 + \sum \alpha_{\eta}^{\zeta} \left[ln(\alpha_{\eta}^{\zeta})\right] \quad (1)$$

where α_{η}^{ζ} represents the frequency of each residue η occurring at position ζ and 5 represents the four possible residues as nucleotides plus gap. Furthermore, maximum entropy per position is taken as 0.2 with no gaps. All these values are taken after following the literature. Thereafter, analysis of the functional characteristics for the mutations in the Spike glycoprotein for the different variants are carried out. Finally, these mutations for each of the variants are visualised in the Spike glycoprotein structure as well.

3. Results

SARS-CoV-2 infects the human cell and after attaching itself to the receptor cell ACE2, it makes the replica of their RNA. Whenever the virus replicates, sometimes the change or mutation is trivial, but whenever the virus changes one or more times it is referred to as a new variant of the original virus. There are several variants that have been reported for SARS-CoV-2. To study these variants in this work, initially multiple sequence alignment of 77681 global SARS-CoV-2 genomic sequences collected from January 2020 to July 2021 is carried out using MAFFT followed by their phylogenetic analysis using Nextstrain. The statistics of the number of sequences considered from each country is reported in Table 1. The phylogenetic analysis of the sequences are given in Fig. 2. After the analysis is completed, in this study, we have reported the 12 important variants or lineages and the corresponding mutations of such variants are reported in Table 2. For example, Alpha first identified in the United Kingdom is characterised by a surprising number of mutations such as H69-, V70-, Y144-L452R, E484K, S494P, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H and K1191N. When compared to the parental strain or the reference sequence, there is a possibility that this variant is associated with a higher viral load and prolonged viral persistence [4] as well as an increased risk of death [3]. Also, epidemiological investigations suggested that Alpha is more transmissible (43–82% higher) than the existing lineages [12]. Beta variant discovered in South Africa [39] has D80A, D215G, L241-, L242-, A243-, P384L, K417N, E484K, N501Y, E516Q, D614G and A701V mutations. This variant has four mutation points K417N, E484K, N501Y and E516Q present in the RBD region of the Spike glycoprotein, thus making it easier for the virus to attach itself to ACE2. Also this variant has been known to significantly reduce neutralisation in antibodies [34]. It also possibly has increased the fatality rate. Preliminary study by Centre of Mathematical Modelling of Infectious Diseases (CMMID COVID-19 working group, London School of Hygiene and Tropical Medicine) has shown that Beta is more transmissible and less susceptible to cross-protection from previous exposure[8]. Epsilon variant was first

[5] https://www.gisaid.org/.
[6] https://www.ncbi.nlm.nih.gov/nuccore/1798174254.
[7] https://zhanglab.ccmb.med.umich.edu/COVID-19/.
[8] https://cmmid.github.io/topics/covid19/sa-novel-variant.html.

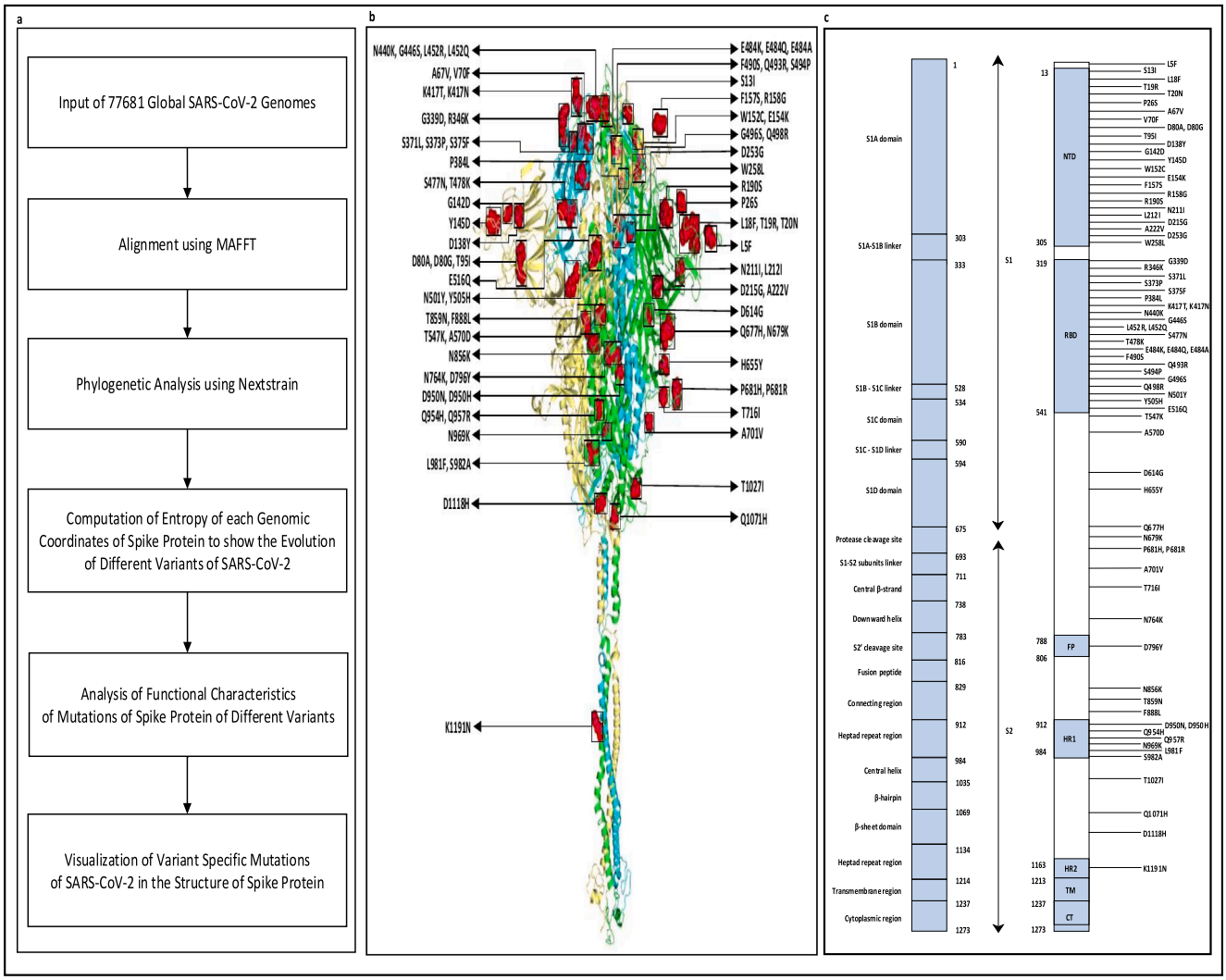

Fig. 1. Pipeline of the Work.

found in USA with the mutation points S13I, W152C, L452R and D614G. In-vitro and epidemiological studies have suggested that this lineage is related to high transmissibility and infectivity. It is also known to escape neutralisation convalescent plasma and antibodies induced by vaccine [12]. Eta variant found in Nigeria has the mutation points A67V, H69-, V70-, Y144-, E484K, D614G, Q677H and F888L. Iota variant found in USA has mutations such as L5F, D80G, T95I, Y144-, F157S, D253G, L452R, S477N, E484K, D614G, A701V, T859N, D950H and Q957R. Discovered in India, Kappa variant has mutations like T95I, G142D, E154K, L452R, E484Q, D614G, P681R and Q1071H. On the other hand, mutations like T19R, V70F, T95I, G142D, E156-, F157-, R158G, A222V, W258L, K417N, L452R, T478K, D614G, P681R and D950N are found in Delta variant which was also discovered in India. Delta variant was responsible for the surge in the number of cases and hospitalisation during the second wave in India. Lambda variant found in Peru has mutations such as L452Q, F490S and D614G. Gamma variant found in Brazil has mutations like L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, P681H and T1027I. It is estimated to be around 2.6 times more transmissible. The efficacy of therapeutic monoclonal antibodies (mAbs) like bamlanivimab, casirivimab and etesivimab may be reduced or even abolished against Gamma. Zeta and Theta variants discovered in Brazil and the Philippines have mutations such as E484K, D614G, N501Y, D614G and P681H. The newly discovered Omicron variant which is currently the dominant variant in most parts of the world has a lot of mutations as compared to the previous variants such as A67V, T95I, G142D, Y145D, N211I, L212I, G339D, R346K, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, Q496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K and L981F. All the mutation details for the different variants along with the entropy values are reported in Table 3. Please note that Omicron shares some mutations (A67V, T95I, G142D, K417N, S477N, T478K, N501Y, D614G and P681H) with other variants like Alpha, Beta, Eta, Iota, Kappa, Delta, Gamma and Theta. Thus, these mutations would have the same entropy as mentioned in Table 3. The rest of the unique mutations pertaining to Omicron should be available for the sequences from November onward and thus their entropies are not very conclusive at the moment. Therefore, they are not included in the analysis hereafter.

The entropy for 77681 SARS-CoV-2 genomes are shown in Fig. 2(c) while the average entropy for each month is visualised in Fig. 3. As can be seen from Fig. 3, the month of March 2020 shows high entropy which even coincides with the 1st wave that swept through the world. Then there was a dip from April to October 2020. During June 2021, again the entropy has a steep rise which marked the 2nd wave. The month wise virus evolution in terms of entropy for the different mutations are visualised in Fig. 4 while the month wise evolution of the mutations pertaining to the different variants like Alpha, Beta, Epsilon, Eta, Iota, Kappa, Delta, Lambda, Gamma, Zeta and Theta are shown in Fig. 5 respectively.

The percentage and frequency of change of nucleotide and amino

Table 1
Statistics of SARS-CoV-2 genomes in different countries.

Name of the Country	Number of Sequences	Name of the Country	Number of Sequences	Name of the Country	Number of Sequences	Name of the Country	Number of Sequences
USA	13387	Northern Ireland	535	Turkey	93	Pakistan	19
England	12126	Luxembourg	530	Peru	90	Hungary	17
India	10307	Canada	496	Slovenia	90	Serbia	16
Scotland	3910	Austria	470	Ghana	82	Belarus	15
Australia	3428	Russia	404	Slovakia	79	Suriname	14
Denmark	2584	Israel	359	Malaysia	79	Georgia	12
Wales	2544	Indonesia	333	Thailand	69	Mali	11
Iceland	1886	Mexico	310	Romania	67	Morocco	11
Belgium	1709	Bangladesh	302	Lithuania	66	Kenya	10
Germany	1690	Norway	267	Croatia	62	Malta	10
Switzerland	1592	Jordan	253	Saudi Arabia	61	Bosnia and Herzegovina	4
Spain	1451	Ecuador	221	Oman	59	Lebanon	4
Netherlands	1432	New Zealand	210	Colombia	53	Bulgaria	4
Italy	1398	Poland	208	North Macedonia	50	Cyprus	4
South Korea	1373	United Arab Emirates	185	Kuwait	45	Guatemala	3
Brazil	1310	Aruba	180	Sri Lanka	44	Kosovo	3
France	1230	Cambodia	169	Argentina	41	Iran	3
Singapore	1127	Greece	151	Curacao	36	Jamaica	3
Japan	976	Latvia	149	Senegal	35	Sierra Leone	3
South Africa	803	Estonia	147	Vietnam	35	Rwanda	2
Sweden	768	Czech Republic	141	Tunisia	31	Brunei	2
China	698	Uganda	130	Costa Rica	30	Panama	1
Finland	669	Egypt	123	Kazakhstan	29	Nepal	1
Portugal	662	Chile	123	Montenegro	25		
Ireland	585	Nigeria	94	Bahrain	23		

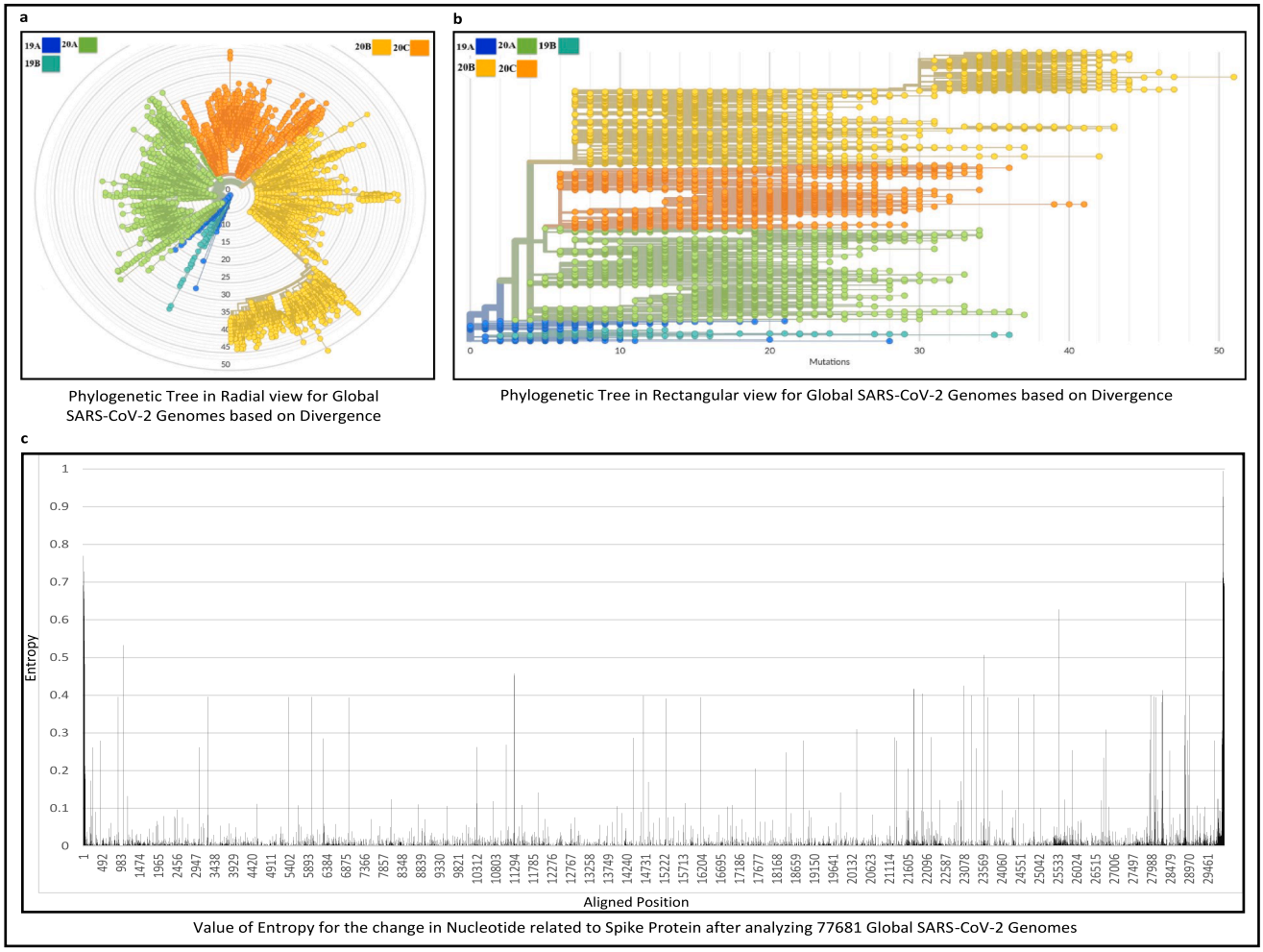

Fig. 2. Phylogenetic analysis of 77681 Global SARS-CoV-2 genomes.

Table 2
Variants of SARS-CoV-2 along with their mutations in Spike Glycoprotein.

Variant (Lineage)	Alpha (B.1.1.7)	Beta (B.1.351)	Epsilon (B.1.427/ B.1.429)	Eta (B.1.525)	Iota (B.1.526)	Kappa (B.1.617.1)	Delta (B.1.617.2)	Lambda (C.37)	Gamma (P.1)	Zeta (P.2)	Theta (P.3)	Omicron (B.1.1.529)
Country of Detection	United Kingdom	South Africa	USA	Nigeria	USA	India	India	Peru	Brazil	Brazil	The Philippines	South Africa
Mutations in Spike Glycoprotein												
L5F					✓							
S13I			✓									
L18F									✓			
T19R							✓					
T20N									✓			
P26S									✓			
A67V				✓								✓
H69-	✓			✓								
V70-	✓			✓								
V70F							✓					
D80A		✓										
D80G					✓							
T95I					✓	✓	✓					✓
D138Y									✓			
G142D						✓	✓					✓
Y144-	✓			✓	✓							
Y145D												✓
W152C			✓									
E154K						✓						
E156-							✓					
F157-							✓					
F157S					✓							
R158G							✓					
R190S									✓			
N211I												✓
L212I												✓
D215G		✓										
A222V							✓					
L241-		✓										
L242-		✓										
A243-		✓										
D253G					✓							
W258L							✓					
G339D												✓
R346K												✓
S371L												✓
S373P												✓
S375F												✓
P384L		✓										
K417T									✓			
K417N		✓					✓					✓
N440K												✓
G446S												✓
L452R	✓		✓		✓	✓	✓					
L452Q								✓				
S477N					✓							✓
T478K							✓					✓
E484A												✓
E484K	✓	✓		✓	✓				✓	✓	✓	
E484Q						✓						
F490S								✓				
Q493R												✓
S494P	✓											
Q496S												✓
Q498R												✓
N501Y	✓	✓							✓		✓	✓
Y505H												✓
E516Q		✓										
T547K												✓
A570D	✓											
D614G	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
H655Y									✓			✓
Q677H				✓								
N679K												✓
P681H	✓								✓		✓	✓
P681R						✓	✓					
A701V		✓			✓							

(*continued on next page*)

Table 2 (*continued*)

Variant (Lineage)	Alpha (B.1.1.7)	Beta (B.1.351)	Epsilon (B.1.427/ B.1.429)	Eta (B.1.525)	Iota (B.1.526)	Kappa (B.1.617.1)	Delta (B.1.617.2)	Lambda (C.37)	Gamma (P.1)	Zeta (P.2)	Theta (P.3)	Omicron (B.1.1.529)
Country of Detection	United Kingdom	South Africa	USA	Nigeria	USA	India	India	Peru	Brazil	Brazil	The Philippines	South Africa
Mutations in Spike Glycoprotein												
T716I	✓											
N764K												✓
D796Y												✓
N856K												✓
T859N					✓							
F888L				✓								
D950N							✓					
D950H					✓							
Q957R					✓							
Q954H												✓
N969K												✓
L981F												✓
S982A	✓											
T1027I									✓			
Q1071H						✓						
D1118H	✓											
K1191N	✓											

Table 3
All mutations in Spike Glycoprotein with relevant details after analysing 77681 Global SARS-CoV-2 genomes.

Mutations in Spike Glycoprotein	Genomic Coordinate	Nucleotide change	Entropy	Mutation in Spike Glycoprotein	Genomic Coordinate	Nucleotide change	Entropy
L5F	21575	C>T	0.1051	L242-	22286	C>-	0.0292
L18F	21614	C>T	0.1917	L242-	22287	T>-	0.0303
S13I	21600	G>T	0.0255	L242-	22288	T>-	0.0279
T19R	21618	C>G	0.2303	A243-	22289	G>-	0.0360
T20N	21621	C>A	0.0976	A243-	22290	C>-	0.0098
P26S	21638	C>T	0.0941	A243-	22291	T>-	0.0102
A67V	21762	C>T	0.0288	D253G	22320	A>G	0.0377
H69-	21767	C>-	0.4524	W258L	22335	G>T	0.0225
H69-	21768	A>-	0.4497	P384L	22713	C>T	0.0115
H69-	21769	T>-	0.4490	K417T	22812	A>C	0.0841
V70F/-	21770	G>T/-	0.4611	K417N	22813	G>T	0.0286
V70-	21771	T>-	0.0401	L452R/Q	22917	T>G/A	0.2774
V70-	21772	C>-	0.0166	S477N	22992	G>A	0.1758
D80A/G	21801	A>C/G	0.0370	T478K	22995	C>A	0.2395
T95I	21846	C>T	0.2267	E484K/Q	23012	G>A/C	0.2041
D138Y	21974	G>T	0.1320	F490S	23031	T>C	0.0180
G142D	21987	G>A	0.3117	S494P	23042	T>C	0.0140
Y144-	21992	T>-	0.4425	N501Y	23063	A>T	0.4805
Y144-	21993	A>-	0.4853	E516Q	23108	G>C	0.0084
Y144-	21994	T>-	0.0713	A570D	23271	C>A	0.4401
W152C	22018	G>T	0.0261	D614G	23403	A>G	0.1576
E154K	22022	G>A	0.0480	H655Y	23525	C>T	0.0905
E156-	22028	G>-	0.0687	Q677H	23593	G>T	0.0659
E156-	22029	A>-	0.2265	P681H/R	23604	C>A/G	0.6381
E156-	22030	G>-	0.2169	A701V	23664	C>T	0.0484
F157-	22031	T>-	0.2167	T716I	23709	C>T	0.4387
F157S/-	22032	T>C/-	0.2410	T859N	24138	C>A	0.0260
F157-	22033	C>-	0.2586	F888L	24224	T>C	0.0089
R158G	22034	A>G	0.2712	D950H/N	24410	G>C/A	0.2490
R190S	22132	G>T	0.0850	Q957R	24432	A>G	0.0238
D215G	22206	A>G	0.0264	S982A	24506	T>G	0.4380
A222V	22227	C>T	0.3203	T1027I	24642	C>T	0.1019
L241-	22283	T>-	0.0261	Q1071H	24775	A>T	0.0475
L241-	22284	T>-	0.0260	D1118H	24914	G>C	0.4439
L241-	22285	A>-	0.0262	K1191N	25135	G>T	0.0307

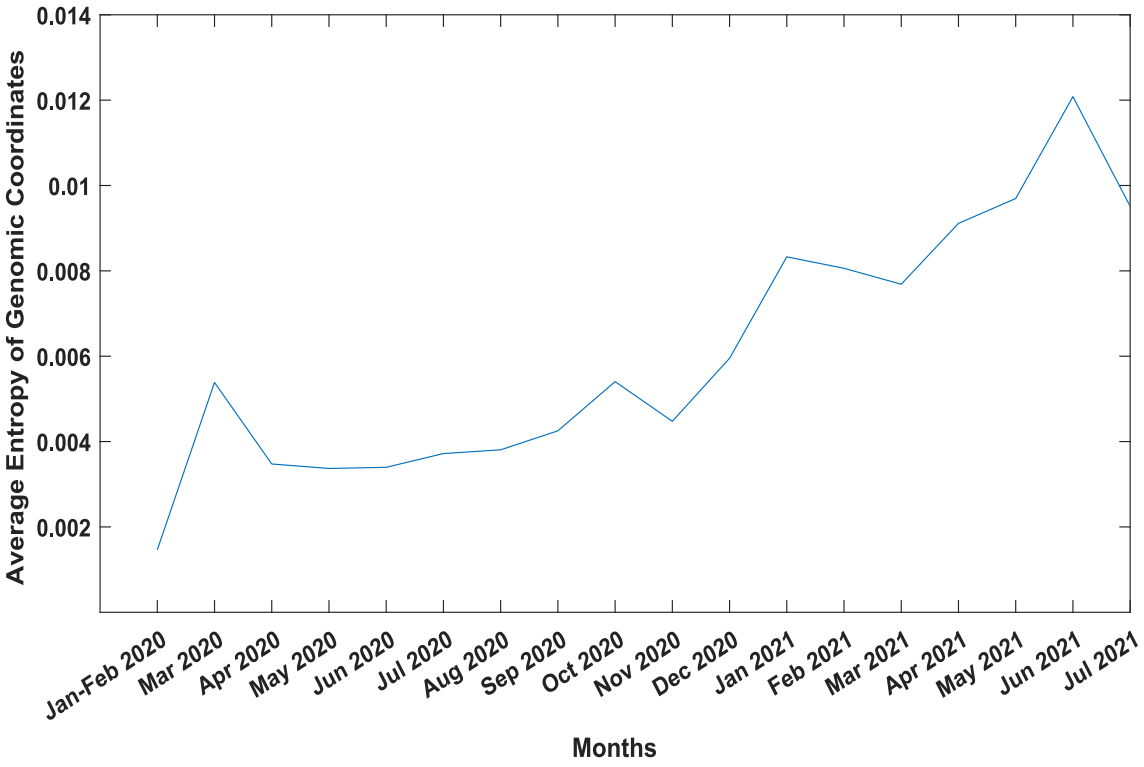

Fig. 3. Average entropy for each month for 77681 Global SARS-CoV-2 genomes.

acid are depicted in Fig. 6 respectively. For example, in Fig. 6(a), the occurrence of T>G in 77681 global SARS-CoV-2 genomes is 18% while Fig. 6(b) shows that the number of times it occurs among 70 nucleotide changes is 2 as is also evident from Table 3. It can also be seen from Fig. 6(b) that 11 out of 70 mutations in Spike glycoprotein are from C to T thereby representing abundant transition. This transition increases the frequency of codons for hydrophobic amino acids and provides evidence of potential anti-viral editing mechanisms driven by host [41]. Also, more C to T transition means less CpG abundance indicating rapid adaptation of virus in host. This CpG deficiency which leads to evasion of host anti-viral defence mechanisms is exhibited the most in SARS-CoV-2 virus [40]. In Fig. 6(c), the occurrence for A>D change in amino acid is 19% while as can be seen from Fig. 6(d), its frequency is 1. All the unique 76 mutations as substitutions corresponding to each of the 12 variant are shown in Fig. 7 along with the structure of Spike glycoprotein.

Structural changes in amino acid residues may sometimes lead to functional instability in proteins due to change in protein translations. These changes are demonstrated through sequence and structural homology-based prediction for the mutations of the different variants in Table 4. Please note that Omicron is not included in this table for the same reason as mentioned before. The tools used for the predictions in Table 4 are PolyPhen-2 (Polymorphism Phenotyping) [1] and I-Mutant 2.0 [5]. Polyphen-2[9] works with sequence, structural and phylogenetic information of mutations while I-Mutant 2.0[10] uses support vector machine (SVM) for the automatic prediction of protein stability changes upon mutations. Polyphen-2 is used to find the damaging mutations and I-Mutant 2.0 determines the corresponding protein stability. To determine if a mutation is damaging using Polyphen-2, its score which lies between 0 and 1 is considered. If the score is close to 1, then a mutation is considered to be damaging. It can be concluded from Table 4 that out of the 53 unique amino acid changes for the 11 variants (apart from Omicron), 22 are damaging. Another important parameter to judge the functional and structural activity of a protein is protein stability which dictates the conformational structure of a protein. Any change in protein stability may cause misfolding, degradation or aberrant conglomeration of proteins. I-Mutant 2.0 uses free energy change values (DDG) to predict the changes in the protein stability wherein a negative value of DDG indicates that the protein has a decreasing stability. The results from I-mutant 2.0 show that out of the 22 unique damaging changes, 18 changes decrease the stability of the protein structures.

4. Discussion

In this section, discussion on the mutation points and the effects of vaccine and therapeutics on the different variants of SARS-CoV-2.

4.1. Characteristics of notable mutation points

There are a total of 84 unique mutation points in the reported 12 SARS-CoV-2 variants. The characteristics of some of the mutations are reported in Table 5.

S13I and W152C are parts of Epsilon variant and help SARS-CoV-2 to escape from therapeutic monoclonal antibodies (mAb). L18F which belongs to Gamma variant helps immune escape from neutralising antibodies (NAbs) against N-terminus. H69- and V70- belonging to Alpha and Eta variants lead to increase in infectivity and reduced sera neutralisation. Y144- present in Alpha, Eta and Iota variants reduce affinity of antibody binding. D253G belonging to Iota variant may aid SARS-CoV-2 to resist NAbs. K417T in Gamma variant is known for resistance to neutralisation by antibodies. The same characteristics is exhibited by K417N which is a part of the Beta and Delta variants. The mutation L452R is part of the Alpha, Epsilon, Iota, Kappa and Delta variants and is largely involved in the significant surge of COVID-19 in India. L452R can increase the binding ability of the ACE2 receptor and can also reduce the attaching capability of vaccine-simulated antibodies with Spike glycoprotein. L452Q belonging to Lambda variant increases viral infectivity. The mutation S477N in Iota and Omicron variants

[9] http://genetics.bwh.harvard.edu/pph2/.
[10] https://folding.biofold.org/i-mutant/i-mutant2.0.html.

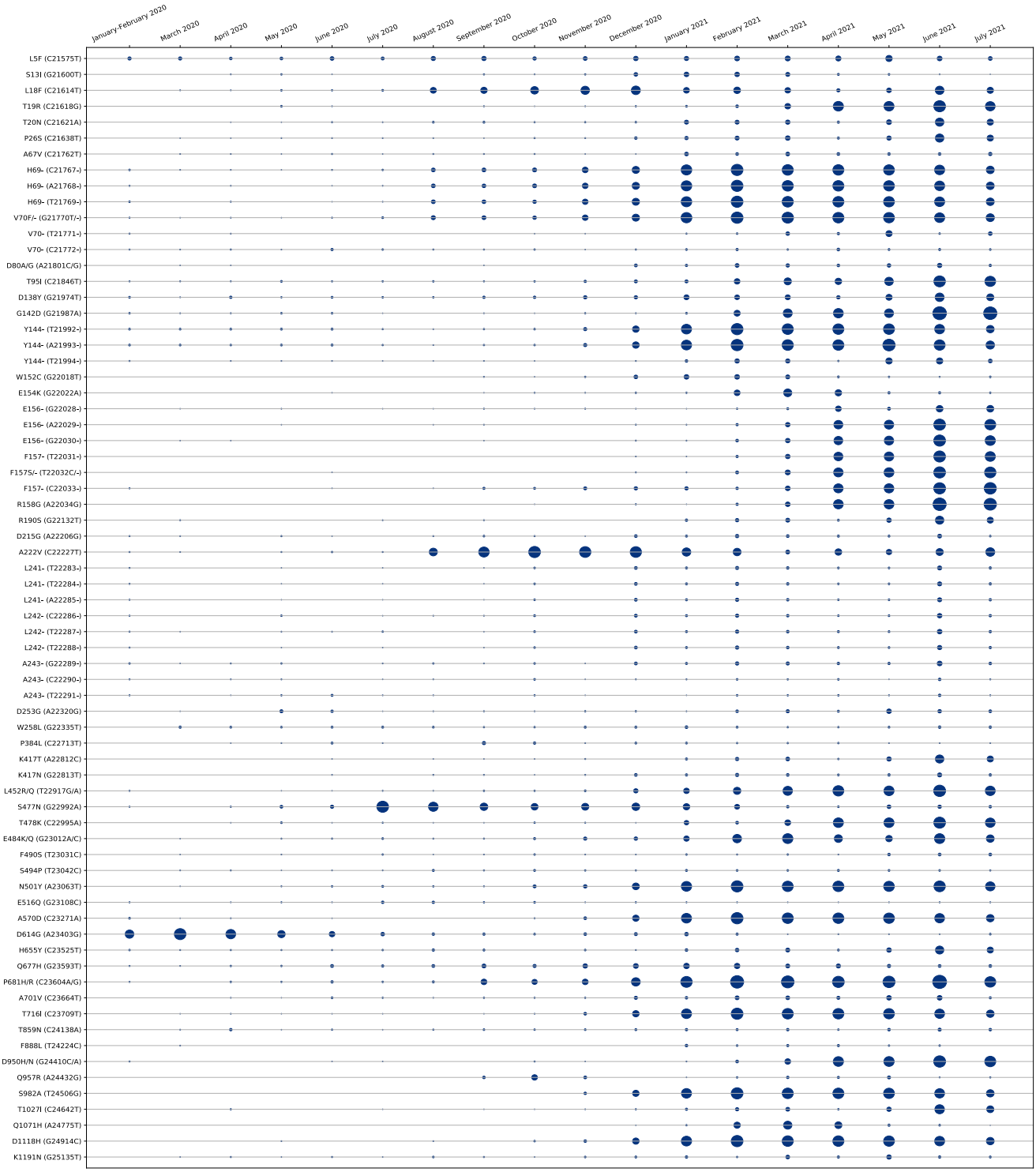

Fig. 4. Month wise evolution of all mutations in Spike Glycoprotein based on entropy after analysing 77681 global SARS-CoV-2 genomes.

present in the RBD region of SARS-CoV-2 results in escape from mAbs. The mutation E484K which is a part of Alpha, Beta, Eta, Iota, Gamma, Zeta and Theta variants is responsible for improving the ability of the virus to escape the host's immune system [17]. Akin to L452R, mutation E484Q also belongs to Kappa variant and is associated with reduced sera neutralisation. F490S in the Lambda variant is associated with reduced susceptibility to antibody neutralisation. The mutation N501Y associated with Alpha, Beta, Gamma, Theta and Omicron variants is present in the receptor binding domain of Spike glycoprotein and has the highest binding affinity with ACE2. N501Y is also known to be associated with immune escape [6]. D614G present in all the 12 reported variants is a significant mutation whose frequency has increased rapidly during the pandemic and is a common mutation in all the lineages or variants. The prevalence of loss of smell has been attributed particularly to this mutation. According to [22], D614G is associated with higher infectivity as well as higher viral load and s1 shedding in Spike glycoprotein. H655Y belonging to Gamma and Omicron variants may affect transmissibility of the virus. Q6777H belonging to Eta variant is also known to affect the

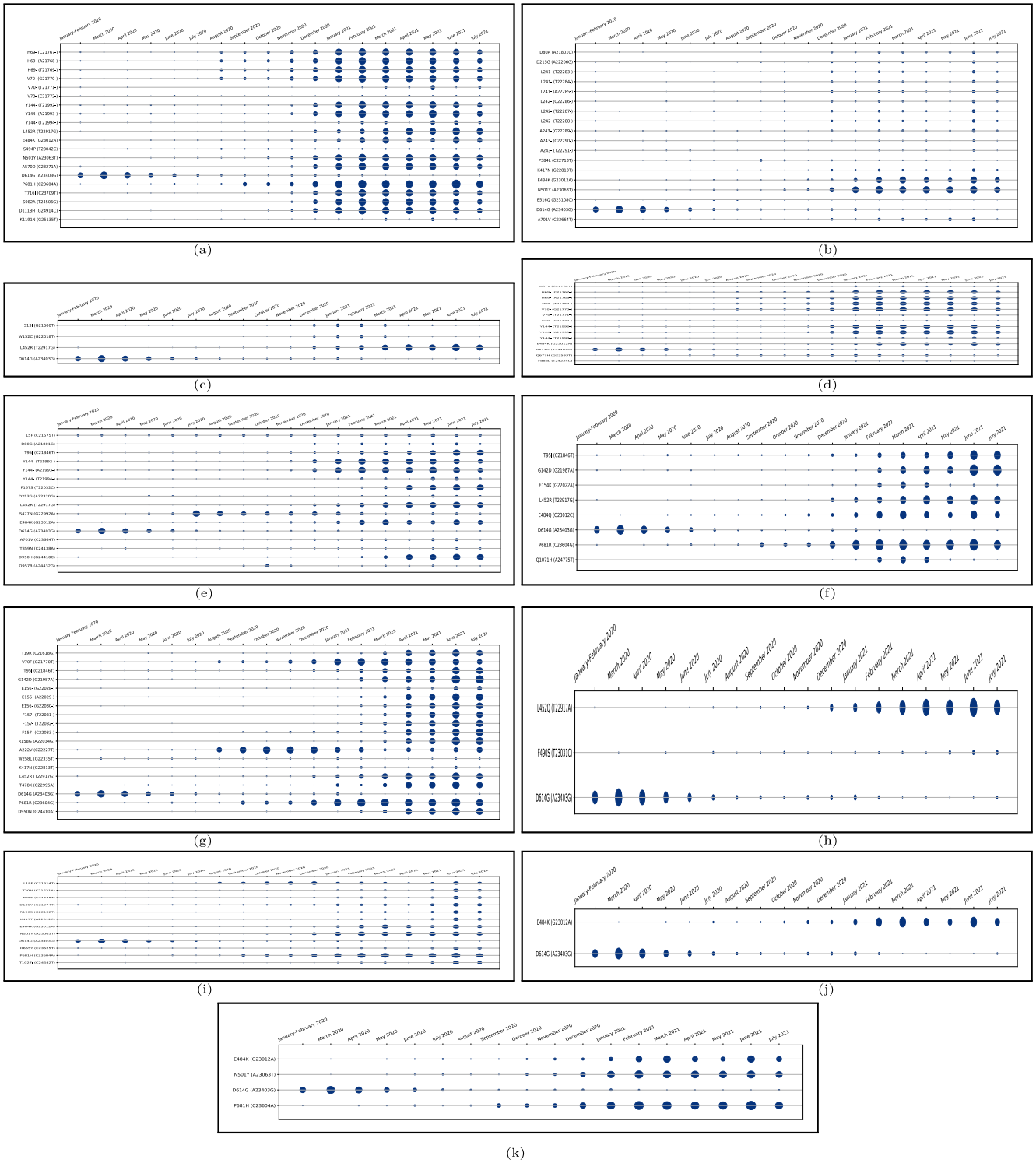

Fig. 5. Month wise evolution of (a) Alpha (B.1.1.7) (b) Beta (B.1.351) (c) Epsilon (B.1.427-B.1.429) (d) Eta (B.1.525) (e) Iota (B.1.526) (f) Kappa (B.1.617.1) (g) Delta (B.1.617.2) (h) Lambda (C.37) (i) Gamma (P.1) (j) Zeta (P.2) and (k) Theta (P.3) variants based on entropy after analysing of 77681 Global SARS-CoV-2 genomes.

transmissibility of SARS-CoV-2. P681H which is a part of Alpha, Gamma and Theta variants and P681R belonging to Kappa and Delta variants have similar functionality as H655Y and Q6777H. In January 2021, scientists reported that similar to D614G, P681H is showing a significant circulation as well and may affect the transmissibility of the virus. Most of the mutations in Omicron like S371L, S373P, S375F, Q493R, and Q498R have high binding affinity with ACE2 receptor. Furthermore, S371L, N440K, G446S and Q493R are also responsible for antibody resistance. It is to be noted that mutations like S371L, S373P, S375F, T478K, Q493R, Q498R and N501Y can induce higher stability in Spike glycoprotein, thereby having high binding affinity with ACE2. This high binding can be attributed to hydrophobic contact at the interfaces of the RBD part of Spike glycoprotein and ACE2 protein [36] and is established by docking studies [23,35] as well.

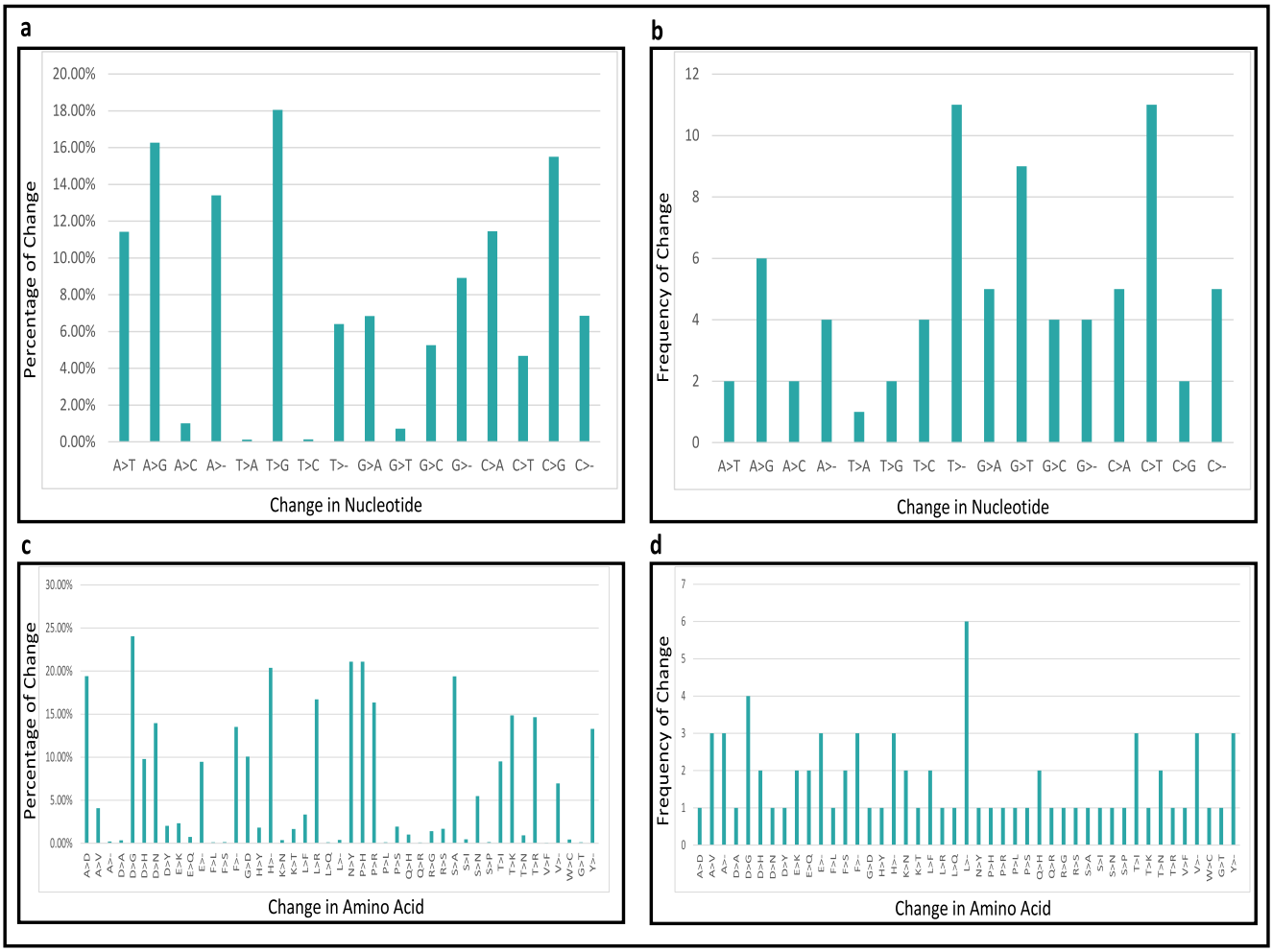

Fig. 6. (a) Percentage of Nucleotide change (b) Frequency of Nucleotide change (c) Percentage of Amino Acid change and (d) Frequency of Amino Acid change for 77681 Global SARS-CoV-2 genomes.

It is to be noted that apart from ACE2, recent research [14] has identified cellular proteins like asialoglycoprotein receptor-1 (ASGR1) and Kringle Containing Transmembrane Protein 1 (KREMEN1) as SARS-CoV-2 receptors in Spike glycoprotein. The authors in [14] have shown that both RBD and N-terminal domain bind of Spike glycoprotein bind to ASGR1 and KREMEN1. These two proteins are also believed to affect the viral target cell range as well as antibody-mediated neutralization [16].

4.2. *Effects of vaccine and therapeutics on different variants*

Vaccines are the most advanced weapon that the human race has devised to fight against this deadly virus. There are several vaccines like Oxford-AstraZeneca, Pfizer-BioNTech, Moderna, Novavax, Covaxin, Sputnik V and Johnson & Johnson which have been developed till now by the scientists around the world. However, some emerging variants like Omicron [26] may be somewhat resistant to the antibody response evoked by these vaccines, thereby making the modifications to these vaccines an absolute necessity. Trials have indicated that many of these vaccines have shown lower efficacy against some of the variants but are effective against the common circulating strains. Table 6 reports the efficacy of the most widely used vaccines for symptomatic as well as severely affected patients. Results have shown that Pfizer-BioNTech and Moderna produced vaccines have an efficacy of 82–100% and 96.3% against the original strain for symptomatic patients while against Delta the efficacy reduces to 42–79% for Pfizer-BioNTech and around 80% for Moderna. For severe patients, efficacy against Delta variant are around 85% and 90% respectively. Gamma variant has been found to partially escape vaccination with Pfizer-BioNTech. Oxford-AstraZeneca vaccine shows an efficacy of 79% against Alpha as opposed to less than 60% against other variants for symptomatic patients. The efficacy of Oxford-AstraZeneca vaccine against Beta was put into question in February 2021 when it was reported that the vaccine is not very effective against this strain. As can be seen from Table 6, the efficacy is indeed very low at 10%. In January 2021, Johnson & Johnson reported that their vaccine was 72% effective against moderate to severe COVID-19 infection in US while such efficiency is 57% in South Africa. According to latest data, Johnson & Johnson vaccine has shown 72% and 86% efficacy in preventing symptomatic COVID-19 and severe COVID-19 respectively for the original strain while for other variants the results vary from 40% to around 75% for both symptomatic and severe patients. Covaxin has also shown promising results for Alpha, Beta, Gamma and Delta variants for symptomatic patients. It is to be noted that Covaxin, Covishield (Indian made Oxford-AstraZeneca vaccine) and Sputnik V have shown effectiveness in neutralising Alpha variant [37]. In March 2021, Novavax vaccine was reported to have a preliminary efficiency of 51% for mild, moderate and severe COVID-19 for HIV-negative patients. According to [11,7], K417N/T, E484K and N501Y are also resistant to neutralisation by vaccines. Despite this, [7] has also reported that sera from infected and Moderna-vaccinated individuals having polyclonal antibodies to the Spike glycoprotein can neutralise the Beta variant. This suggests that protective humoral immunity may be retained against Beta. Research regarding effectiveness of the existing vaccines against the latest

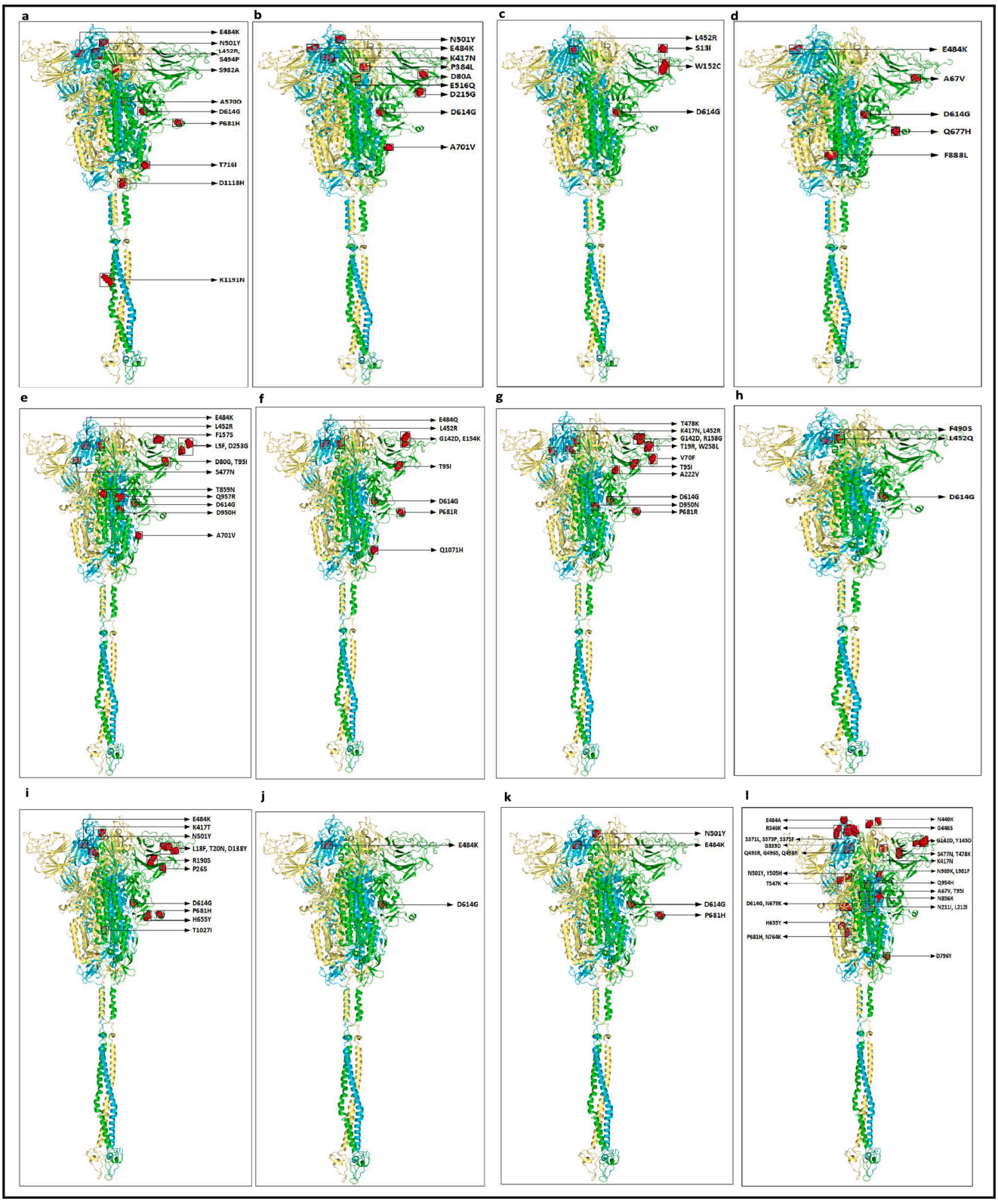

Fig. 7. Highlighted amino acid changes in the Spike glycoprotein of SARS-CoV-2 variants for (a) Alpha (B.1.1.7) (b) Beta (B.1.351) (c) Epsilon (B.1.427-B.1.429) (d) Eta (B.1.525) (e) Iota (B.1.526) (f) Kappa (B.1.617.1) (g) Delta (B.1.617.2) (h) Lambda (C.37) (i) Gamma (P.1) (j) Zeta (P.2) (k) Theta (P.3) and (l) Omicron (B.1.1.529).

circulating Omicron variant is ongoing.

Monoclonal antibody therapies like LY-CoV555 (Bamlanivimab) has been shown to work against Alpha but Beta, Gamma and Epsilon are resistant against it while Alpha, Beta and Gamma variants are resistant against Etesevimab but there is no data for Epsilon variant. Though, Alpha is susceptible to both REGN10933 (Casirivimab) and REGN10987 (Imdevimab), Beta and Gamma are both partially resistant to Casirimivab but Imdevimab is effective against them. As of 22nd December

Table 4
Biological functionality and protein structural stability of the mutations for different variants.

Change in Nucleotide	Change in Amino Acid	PolyPhen-2		I-Mutant 2.0	
		Prediction	Score	Stability	DDG
C21575T	L5F	Not Generated	Not Generated	Decrease	-0.10
G21600T	S13I	Not Generated	Not Generated	Increase	0.39
C21614T	L18F	Possibly Damaging	0.500	Decrease	-0.39
C21618G	T19R	Benign	0.004	Decrease	-0.12
C21621A	T20N	Benign	0.000	Decrease	-0.78
C21638T	P26S	Benign	0.009	Decrease	-2.19
C21762T	A67V	Benign	0.054	Decrease	-0.02
G21770T	V70F	Benign	0.111	Decrease	-2.72
A21801C	D80A	Possibly Damaging	0.858	Decrease	-1.91
A21801G	D80G	Benign	0.016	Decrease	-1.81
C21846T	T95I	Probably Damaging	0.999	Decrease	-1.80
G21974T	D138Y	Probably Damaging	0.992	Increase	1.47
G21987A	G142D	Benign	0.051	Decrease	-1.17
G22018T	W152C	Probably Damaging	0.996	Decrease	-1.66
G22022A	E154K	Not Generated	Not Generated	Decrease	-1.40
T22032C	F157S	Not Generated	Not Generated	Decrease	-2.57
A22034G	R158G	Not Generated	Not Generated	Decrease	-2.63
G22132T	R190S	Probably Damaging	0.996	Decrease	-2.09
A22206G	D215G	Benign	0.002	Decrease	-1.06
C22227T	A222V	Benign	0.001	Increase	0.48
A22320G	D253G	Not Generated	Not Generated	Decrease	-2.43
G22335T	W258L	Benign	0.055	Decrease	-0.61
C22713T	P384L	Probably Damaging	0.972	Decrease	-1.74
A22812C	K417T	Benign	0.012	Decrease	-0.88
G22813T	K417N	Benign	0.341	Decrease	-0.33
T22917G	L452R	Benign	0.040	Decrease	-1.40
T22917A	L452Q	Benign	0.077	Decrease	-1.52
G22992A	S477N	Benign	0.007	Increase	0.01
C22995A	T478K	Benign	0.000	Decrease	-0.09
G23012A	E484K	Benign	0.427	Decrease	-0.85
G23012C	E484Q	Possibly Damaging	0.786	Decrease	-0.48
T23031C	F490S	Benign	0.012	Decrease	-2.99
T23042C	S494P	Possibly Damaging	0.889	Decrease	-0.66
A23063T	N501Y	Benign	0.145	Decrease	-0.34
G23108C	E516Q	Probably Damaging	0.997	Decrease	-0.93
C23271A	A570D	Benign	0.031	Decrease	-1.32
A23403G	D614G	Benign	0.002	Decrease	-1.94
C23525T	H655Y	Benign	0.002	Increase	0.43
G23593T	Q677H	Benign	0.157	Increase	0.10
C23604A	P681H	Not Generated	Not Generated	Decrease	-0.92
C23604G	P681R	Not Generated	Not Generated	Decrease	-0.79
C23664T	A701V	Possibly Damaging	0.887	Increase	0.05
C23709T	T716I	Possibly Damaging	0.696	Decrease	-0.95
C24138A	T859N	Probably Damaging	0.989	Decrease	-0.82
T24224C	F888L	Probably Damaging	0.989	Increase	0.13
G24410A	D950N	Possibly Damaging	0.731	Increase	0.15
G24410C	D950H	Probably Damaging	0.999	Decrease	-0.10
A24432G	Q957R	Possibly Damaging	0.679	Decrease	-0.93
T24506G	S982A	Probably Damaging	0.996	Decrease	-1.36
C24642T	T1027I	Probably Damaging	1.000	Decrease	-0.22
A24775T	Q1071H	Probably Damaging	0.998	Decrease	-1.19
G24914C	D1118H	Probably Damaging	0.998	Decrease	-0.10
G25135T	K1191N	Probably Damaging	0.996	Decrease	-1.40

Table 5
Characteristics of mutations in Spike Glycoprotein.

Mutations	Characteristics
S13I	Helps SARS-CoV-2 to escape from mAbs [30]
L18F	Immune escape from NAbs against N-terminus [31]
H69-	Increase in infectivity and reduced sera neutralisation [32,20]
V70-	Increase in infectivity and reduced sera neutralisation [32,20]
Y144-	Reduces affinity of antibody binding [32]
W152C	Helps SARS-CoV-2 to escape from mAbs [30]
D253G	May aid resistance to neutralising Abs [25]
S371L	High binding affinity with ACE2 [23] and responsible for antibody resistance [26]
S373P	High binding affinity with ACE2 [23]
S375F	High binding affinity with ACE2 [23]
K417T	Resistant to neutralisation [13]
K417N	Resistant to neutralisation [13]
N440K	Responsible for antibody resistance [26]
G446S	Responsible for antibody resistance [26]
L452R	Increases the binding ability of the ACE2 receptor and can also reduce the attaching capability to vaccine [10]
L452Q	Increases viral infectivity [21]
S477N	Results in escape from mAbs [27]
T478K	High binding affinity with ACE2 [23]
E484K	Responsible for improving the ability of the virus to escape the host's immune system [18]
E484Q	Associated with reduced sera neutralisation [13]
F490S	Associated with reduced susceptibility to antibody neutralization [21]
Q493R	High binding affinity with ACE2 [23] and responsible for antibody resistance [26]
Q498R	High binding affinity with ACE2 [23]
N501Y	Highest binding affinity with ACE2 and resistant to neutralisation [28]
D614G	Associated with higher infectivity as well as higher viral load and s1 shedding in Spike glycoprotein [22]
H655Y	Near furin cleavage site, may affect transmissibility of the virus[10]
Q677H	Near furin cleavage site, may affect transmissibility of the virus [2]
P681H	Near furin cleavage site, may affect transmissibility of the virus [2]
P681R	Near furin cleavage site, may affect transmissibility of the virus [2]

[10] https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-omicron-variant.html.

2021, FDA has authorised Pfizer's Paxlovid for the treatment of mild-to-moderate COVID-19 disease in adults and pediatric patients.

5. Conclusion

In this work, we have provided a comprehensive study of the different important variants of SARS-CoV-2 and their corresponding unique mutation points in Spike glycoprotein. This is especially important to understand the effect of the mutations on the vaccines. In this regard, there are 12 important variants of SARS-CoV-2 which are identified; they being Alpha, Beta, Eta, Epsilon, Iota, Kappa, Delta, Lambda, Gamma, Zeta, Theta and lately, Omicron and they have 84 unique mutations in the Spike glycoprotein. These 84 include such mutations like S371L, N440K, G446S, Q493R, N501Y etc. which are

Table 6
Efficacy of vaccines against different variants of SARS-CoV-2.

Vaccine	Symptomatic					Severe				
	Original Virus	Alpha	Beta	Gamma	Delta	Original Virus	Alpha	Beta	Gamma	Delta
BNT162b2 (Pfizer-BioNTech)	82–100%[11] [9]	78–95% [9]	75% [9]	No published data	42–79% [9]	75–95% [9]	>95% [29]	>95% [29]	95% [9]	>85% [29]
mRNA-1273 (Moderna)	96.3% [33]	84–99% [9]	>80% [29]	>95% [29]	>80% [29]	No published data	>90% [29]	No published data	No published data	>90% [29]
AZD1222 (Oxford Astrazeneca)	76%[11] 85% for people over 60[11]	79% [9]	10% [29]	>60% [29]	>60% [29]	>80% [29]	No published data	No published data	No published data	>80% [29]
Janssen (Johnson & Johnson)	72%[11]	>75% [29]	~40% [29]	~ 40% [29]	47–79% [9]	86%[11]	No published data	72% [2]	>60% [29]	>60% [29]
Covaxin(Bharat Biotech)	77.8% [8] 68% for people over 60 [8]	71%[12]	71%[12]	71%[12]	65.2% [8]	93.4% [8]	No published data	No published data	No published data	No published data

[11] https://www.yalemedicine.org/news/covid-19-vaccine-comparison.
[12] https://www.who.int/news-room/feature-stories/detail/the-bharat-biotech-bbv152-covaxin-vaccine-against-covid-19-what-you-need-to-know.

known to resist antibodies. With the current surge of Omicron variant throughout the world and it being highly resistant to neutralisation by the existing vaccines, booster shots are being recommended worldwide and new phases of partial lockdowns are also coming into effect. In this current scenario, the existing vaccines are getting modified and new vaccines are also being manufactured. We hope that this work provides the readers a comprehensive review of the emerging variants and the characteristics of the corresponding mutation points along with the effects of vaccine and therapeutics on the variants.

Ethics approval and consent to participate

The ethical approval or individual consent was not applicable.

Availability of data and materials

The aligned 77681 SARS-CoV-2 genomes with reference sequence are available at "http://www.nitttrkol.ac.in/indrajit/projects/COVID-SpikeVariantsReview-77K".

Consent for publication

Not applicable.

Funding

This work was carried out during the tenure of an ERCIM 'Alain Bensoussan' Fellowship Program awarded to Dr. Nimisha Ghosh. This work has also been partially supported by CRG short term research grant on COVID-19 (CVD/2020/000991) from Science and Engineering Research Board (SERB), Department of Science and Technology, Govt. of India..

Author contributions

Nimisha Ghosh: Conceptualization; Data curation; Formal analysis; Validation; Visualization; Writing - original draft, **Suman Nandi**: Conceptualization; Formal analysis; Software; Validation; Visualization; Writing - review and editing, **Indrajit Saha**: Conceptualization; Data curation; Supervision; Formal analysis; Investigation; Project administration; Resources; Validation; Writing - review and editing.

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.

Acknowledgment

The authors thank all those who have contributed sequences to GISAID database.

References

[1] I.A. Adzhubei, S. Schmidt, L. Peshkin, et al., A method and server for predicting damaging missense mutations, Nat. Methods 7 (2010) 248–249, https://doi.org/10.1038/nmeth0410-248.
[2] E. Boehm, I. Kronig, R.A. Neher, et al., Novel SARS-CoV-2 variants: the pandemics within the pandemic, Clin. Microbiol. Infect. 27 (8) (2021) 1109–1117, https://doi.org/10.1016/j.cmi.2021.05.022.
[3] T. Burki, Understanding variants of sars-cov-2, The Lancet 397 (2021) 462, https://doi.org/10.1016/S0140-6736(21)00298-1.
[4] P. Calistri, L. Amato, I. Puglia, et al., Infection sustained by lineage B.1.1.7 of SARS-CoV-2 is characterised by longer persistence and higher viral RNA loads in nasopharyngeal swabs, Int. J. Infect. Dis. 105 (2021) 753–755, https://doi.org/10.1016/j.ijid.2021.03.005.
[5] E. Capriotti, P. Fariselli, R. Casadio, I-mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure, Nucleic Acids Res. 33 (2005) 306–310, https://doi.org/10.1093/nar/gki375.
[6] P. Colson, A. Levasseur, J. Delerce, et al., Spreading of a new SARS-CoV-2 N501Y spike variant in a new lineage, Clin. Microbiol. Infect. (2021), https://doi.org/10.1016/j.cmi.2021.05.006.
[7] V.V. Edara, C. Norwood, K. Floyd, et al., Infection- and vaccine-induced antibody binding and neutralization of the B.1.351 SARS-CoV-2 variant, Cell Host Microbe 29 (4) (2021) 516–521.e3, https://doi.org/10.1016/j.chom.2021.03.009.
[8] R. Ella, S. Reddy, W. Blackwelder, et al., Efficacy, safety, and lot-to-lot immunogenicity of an inactivated SARS-CoV-2 vaccine (BBV152): interim results of a randomised, double-blind, controlled, phase 3 trial, The Lancet 398 (2021) 2173–2184, https://doi.org/10.1016/S0140-6736(21)02000-6.
[9] T. Fiolet, Y. Kherabi, C.J. MacDonald, et al., Comparing COVID-19 vaccines for their characteristics, efficacy and effectiveness against SARS-CoV-2 and variants of concern: a narrative review, Clin. Microbiol. Infect. (2021), https://doi.org/10.1016/j.cmi.2021.10.005.
[10] W.F. Garcia-Beltran, E.C. Lam, K.S. Denis, et al., Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity, Cell 184 (9) (2021) 2372–2383.e9, https://doi.org/10.1016/j.cell.2021.03.013.
[11] W.F. Garcia-Beltran, E.C. Lam, K. St. Denis, et al., Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity, Cell 184 (9) (2021) 2372–2383.e9, https://doi.org/10.1016/j.cell.2021.03.013.

[12] F. González-Candelas, M.A. Shaw, T. Phan, et al., One year into the pandemic: Short-term evolution of SARS-CoV-2 and emergence of new lineages, Infect. Genet. Evol. 92 (2021) 104869, https://doi.org/10.1016/j.meegid.2021.104869.
[13] A. Greaney, T. Starr, P. Gilchuk, et al., Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition, Cell Host Microbe 29 (1) (2021) 44–57.e9, https://doi.org/10.1016/j.chom.2020.11.007.
[14] Y. Gu, J. Cao, X. Zhang, et al., Receptome profiling identifies KREMEN1 and ASGR1 as alternative functional receptors of SARS-CoV-2, Cell Res. 32 (2021) 24–37, https://doi.org/10.1038/s41422-021-00595-6.
[15] L. Guruprasad, Human SARS-CoV-2 spike protein mutations, Proteins: Struct., Funct., Bioinf. 89 (2021) 569–576, https://doi.org/10.1002/prot.26042.
[16] M. Hoffmann, S. Pöhlmann, Novel SARS-CoV-2 receptors: ASGR1 and KREMEN1, Cell Res. 32 (2021) 1–2, https://doi.org/10.1038/s41422-021-00595-6.
[17] S. Jangra, C. Ye, R. Rathnasinghe, et al., SARS-CoV-2 spike E484K mutation reduces antibody neutralisation, Lancet Microbe (2021), https://doi.org/10.1016/S2666-5247(21)00068-9.
[18] S. Jangra, C. Ye, R. Rathnasinghe, et al., SARS-CoV-2 spike E484K mutation reduces antibody neutralisation, The Lancet 2 (7) (2021) E283–E284, https://doi.org/10.1016/S2666-5247(21)00068-9.
[19] K. Katoh, K. Misawa, K. i Kuma, et al., MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, Nucleic Acids Res. 30 (14) (2002) 3059–3066, https://doi.org/10.1093/nar/gkf436.
[20] S.A. Kemp, B. Meng, I.A. Ferriera, et al., Recurrent emergence and transmission of a sars-cov-2 spike deletion h69/v70. BioRxiv, 2021.
[21] I. Kimura, Y. Kosugi, J. Wu, et al., The SARS-CoV-2 Lambda variant exhibits enhanced infectivity and immune resistance, Cell Reports 38 (2) (2022) 110218, https://doi.org/10.1016/j.celrep.2021.110218.
[22] B. Korber, W.M. Fischer, S. Gnanakaran, et al., Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus, Cell 182 (2020) 812–827, https://doi.org/10.1016/j.cell.2020.06.043.
[23] S. Kumar, T.S. Thambiraja, K. Karuppanan, et al., Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein, J. Med. Virol. (2021), https://doi.org/10.1002/jmv.27526.
[24] S.D. Lam, N. Bordin, V.P. Waman, et al., SARS-CoV-2 spike protein predicted to form complexes with host receptor protein orthologues from a broad range of mammals, Sci. Rep. 10 (2020) 1–14, https://doi.org/10.1038/s41598-020-71936-5.
[25] E. Lasek-Nesselquist, P. Lapierre, E. Schneider et al., The localized rise of a B. 1.526 variant containing an E484K mutation in New York State. medRxiv, 2021.
[26] L. Liu, S. Iketani, Y. Guo, et al., Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2, Nature (2021), https://doi.org/10.1002/jmv.27526.
[27] Z. Liu, L.A. VanBlargan, L.M. Bloyet, et al., Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization, Cell Host Microbe 29 (3) (2021) 477–488.e4, https://doi.org/10.1016/j.chom.2021.01.014.
[28] B. Luan, H. Wang, T. Huynh, 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) (2021) 1454–1461, https://doi.org/10.1002/1873-3468.14076.
[29] S. Mallapaty, E. Callaway, M. Kozlov, et al., How COVID vaccines shaped 2021 in eight powerful charts, Nature 600 (2021) 580–583, https://doi.org/10.1038/d41586-021-03686-x.
[30] M. McCallum, J. Bassi, A.D. Marco, et al., SARS-CoV-2 immune evasion by variant B. 1.427/B. 1.429. BioRxiv, 2021.
[31] M. McCallum, A.D. Marco, F.A. Lempp, et al., N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2, Cell 184 (9) (2021) 2332–2347, https://doi.org/10.1016/j.cell.2021.03.028.
[32] K.R. McCarthy, L.J. Rennick, S. Nambulli, et al., Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape, Science 371 (6534) (2021) 1139–1142, https://doi.org/10.1126/science.abf6950.
[33] T. Pilishvili, R. Gierke, K. Fleming-Dutra, et al., Effectiveness of mRNA Covid-19 vaccine among U.S. health care personnel, N. Engl. J. Med. 385 (25) (2021) e90, https://doi.org/10.1056/NEJMoa2106599.
[34] D. Planas, T. Bruel, Grzelak, et al., Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies, Nature Med. 27 (2021) 917–924, https://doi.org/10.1038/s41591-021-01318-5.
[35] I. Saha, N. Ghosh, N. Sharma, et al., Hotspot mutations in SARS-CoV-2, Front. Genetics 12 (2021) 2076, https://doi.org/10.3389/fgene.2021.753440.
[36] S.Y. Sathipati, S.K. Shukla, S.Y. Ho, Tracking the amino acid changes of spike proteins across diverse host species of severe acute respiratory syndrome coronavirus 2, iScience 25 (1) (2022) 103560, https://doi.org/10.1016/j.isci.2021.103560.
[37] J. Singh, S.A. Rahman, N.Z. Ehtesham, et al., SARS-CoV-2 variants of concern are emerging in India, Nat. Med. (2021), https://doi.org/10.1038/s41591-021-01397-4.
[38] C.W. Tan, W.N. Chia, X. Qin, et al., A SARS-CoV-2 surrogate virus neutralization test based on antibody-mediated blockage of ACE2–spike protein–protein interaction, Nature Biotechnol. 38 (9) (2020) 1073–1078, https://doi.org/10.1038/s41587-020-0631-z.
[39] H. Tegally, E. Wilkinson, M. Giovanetti, et al., Detection of a SARS-CoV-2 variant of concern in South Africa, Nature 592 (2021) 438–443, https://doi.org/10.1038/s41586-021-03402-9.
[40] X. Xia, Extreme Genomic CpG Deficiency in SARS-CoV-2 and Evasion of Host Antiviral Defense, Mol. Biol. Evol. 37 (9) (2020) 2699–2705, https://doi.org/10.1093/molbev/msaa094.
[41] F. Yuan, L. Wang, Y. Fang, et al., Global SNP analysis of 11,183 SARS-CoV-2 strains reveals high genetic diversity, Transboundary Emerg. Dis. (2020), https://doi.org/10.1111/tbed.13931.
[42] L. Zhang, C.B. Jackson, H. Mou, et al., SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity, Nature Commun. 11 (2020) 1–9, https://doi.org/10.1038/s41467-020-19808-4.