

Research Article

Annotating Spike Protein Polymorphic Amino Acids of Variants of SARS-CoV-2, Including Omicron

Gusti Ngurah Mahardika ¹, Nyoman B. Mahendra ², Bayu K. Mahardika ³,
Ida B. K. Suardana ⁴ and Made Pharmawati ⁵

¹The Animal Biomedical and Molecular Biology Laboratory, Udayana University, Jl. Sesetan-Markisa 6A, Denpasar 80223, Bali, Indonesia

²The Department of Obstetrics and Gynecology, The Faculty of Medicine, Udayana University, Denpasar, Bali, Indonesia

³The Animal Biomedical and Molecular Biology Laboratory, Udayana University, Jl. Sesetan-Markisa 6A, Denpasar 80223, Bali, Indonesia

⁴Virology Laboratory, The Faculty of Veterinary Medicine, Udayana University, Denpasar, Bali, Indonesia

⁵The Biology Study Program, The Faculty of Mathematic and Natural Science, Udayana University, Kampus Bukit Jimbaran, Badung, Bali, Indonesia

Correspondence should be addressed to Gusti Ngurah Mahardika; gnmahardika@unud.ac.id

Received 30 January 2022; Accepted 24 March 2022; Published 9 April 2022

Academic Editor: Saleh Ahmed Mohamed

Copyright © 2022 Gusti Ngurah Mahardika et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The prolonged global spread and community transmission of severe acute respiratory syndrome virus 2 (SARS-CoV-2) has led to the emergence of variants and brought questions regarding disease severity and vaccine effectiveness. We conducted simple bioinformatics on the spike gene of a representative of each variant. The data show that a number of polymorphic amino acids are located mostly on the amino-terminal side of the S1/S2 cleavage site. The Omicron variant diverges from the others, with the highest number of amino acid substitutions, including the receptor-binding site (RBS), epitopes, S1/S2 cleavage site, fusion peptide, and heptad repeat 1. The current sharp global increase in the frequency of the Omicron genome constitutes evidence of its high community transmissibility. In conclusion, the proposed guideline could give an immediate insight of the probable biological nature of any variant of SARS-Cov-2. As the Omicron diverged the farthest from the original pandemic strain, Wuhan-Hu-1, we expect different epidemiological and clinical patterns of Omicron cases. On vaccine efficacy, slight changes in some epitopes while others are conserved should not lead to a significant reduction in the effectiveness of an approved vaccine.

1. Introduction

The emergence of various variants of the severe acute respiratory syndrome virus 2 (SARS-CoV-2) has led to questions about disease severity and vaccine effectiveness. As of 19 December 2021, over 273 million cases and over 5.3 million deaths have been reported globally (<https://www.who.int>). The World Health Organization defines SARS-CoV-2 variants as variants of concern (VOCs), variants of interest (VOIs), and variants under monitoring (VUMs). The VOCs listed in the GISAID database, which was accessed on December 26, 2021, are Omicron,

Delta, Alpha, Beta, and Gamma; Lambda and Mu are VOIs; and GH/490R is a VUM.

The impact of amino acid substitutions in the variants must depend on their genetic make-up. The SARS-CoV-2 belongs to the lineage B (sarbecovirus) of β -CoVs of the *Coronaviridae* family, which is enveloped with single-stranded positive-sense RNA [1]. The genome of SARS-CoV-2 is more than 29 kb, which encodes for structural proteins of phosphorylated nucleocapsid (N) protein, spike glycoprotein (S), hemagglutinin-esterase (HE), membrane (M) protein, and envelope (E) protein, as well as nonstructural proteins of ORF1a, ORF1b, ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8, ORF9a, ORF9b,

and ORF10 [1, 2]. ORF1a and 1b are also translated following -1 ribosomal frameshifting to produce ORF1ab protein [3]. Proteins translated from ORF1a, 1b, and 1ab form viral polymerase complex [4], while proteins of other ORF's are also known as accessory proteins [5]. However, it is generally believed that the spike protein is a major pathogenic coronavirus determinant. This surface protein possesses major immunogenic domains, and most gene-based vaccines target only the spike gene of SARS-CoV-2. The protein is highly glycosylated and cysteine rich, with two cleavage sites: S1/S2 and S2' [6]. The glycosylation pattern of the SARS-CoV-2 spike involves N-linked and O-linked glycosylation [7]. Spike also has two protease cleavage sites, which are critical for virus activation and replication. Protease cleavage of the spike has been established as a critical determinant of coronavirus tropism and pathogenesis [8].

The epitopes of the spike protein seem to be linear and conformational. One group [9] mapped nine linear epitopes along the spike protein designated IdA-IdI; another group [10] identified 16 epitopes. Some epitopes overlap. Due to folding and trimerization, conformational epitopes in the spike protein of SARS-CoV-2 have been predicted [11].

The biological consequences of each variant are poorly understood. In this study, we performed functional annotation of amino acid changes in the spike protein of variants based on existing knowledge of coronaviruses as well as rapidly accumulating knowledge about SARS-CoV-2.

2. Materials and Methods

The sequence of the original SARS-CoV-2 strain of Wuhan-Hu-1 (Accession Number NC_045512) was downloaded from GenBank. The open reading frame (ORF) of the spike protein was selected as determined in the database. Ten complete sequences of each annotated variant were selected randomly from GISAID and downloaded. Using the spike gene of Wuhan-Hu-1, the first 15 nucleotides of the 5'-terminus were searched, and the sequence prior to the marked sequence was deleted. The last 15 nucleotides of Wuhan-Hu-1 were used to mark the 3'-end of the sequence, and all nucleotides after that marked sequence were deleted. We then manually selected the sequence data without any undefined or any "N" residue. When there was no "clean" sequence data for each variant, we evaluated another set of variant data. The selected sequences were translated into amino acid sequences and aligned using MEGA-X software [12]. Using the same software, the data were exported in Mega format and analyzed further for polymorphic or variable amino acids. The evolutionary history of variants was inferred using the neighbor-joining method [13]. Evolutionary distances were computed using the Kimura 2-parameter method [14]. The probable biological function of each residue was annotated using the guidelines shown in Supplementary Material 1.

3. Results

The dataset containing the representative of each variant used is available in Supplementary Material 2. Polymorphic amino acid residues of the spike protein of SARS-CoV-2

Wuhan Hu-1 and all variants are presented in Table 1. The data show that a number of polymorphic amino acids are located in the S1 domain, on the N-terminal side of the S1/S2 cleavage site. The number of polymorphic amino acids in this region is 73; the S2 domain has only 15. Glycosylation motif loss (GML) occurs once in Delta due to the T19R substitution and in Lambda due to the T76I substitution. Additional glycosylation motif (AGM) gain occurred twice in the Gamma variant, i.e., T20N and R191S, and once in Lambda, i.e., R249N. Cysteine residue loss (CRL) occurred once in the GH/490 variant due to the deletion of C136. Ten residues of the receptor binding site (RBS) are polymorphic, with only a single residue difference from Wuhan-Hu-1 in the Alpha and GH variants; the other variants, except for Omicron, carry two residues, and Omicron shows nine amino acid differences from Wuhan-Hu-1. The number of amino acid substitutions in various linear epitopes is 19; that of probable conformational epitopes is 16. The number of amino acid changes from Wuhan-Hu-1 at the mapped linear epitopes of various variants is 2, 6, 4, 2, 3, 3, 10, and 3 for Alpha, Beta, Gamma, Delta, Lambda, Mu, Omicron, and GH/490R, respectively. A single amino acid difference from Wuhan-Hu-1 at probable conformational epitopes occurred in Beta, Gamma, Delta, Mu, and GH/490R; three changes in Alpha, and seven each occurred in Lambda and Omicron. At the S1/S2 cleavage site of the Alpha, Beta, Delta, Lambda, and GH/490R variants differ from Wuhan-Hu-1 in one residue, with Omicron differing in two. All residues at this site have changed from the nonbasic amino acids Q/N/P to the basic amino acids H/R/K. At the fusion peptide site, only a single amino acid substitution occurred in Omicron. In heptad repeat 1 (HR1), a single amino acid alteration occurred in the Alpha, Gamma, and Mu variants, with Omicron displaying three alterations. In heptad repeat 2 (HR2) and the transmembrane domain, a single amino acid change occurred in the Gamma variant only.

The phylogenetic analysis presented in Figure 1 shows two clusters of variants, with good bootstrap support of 88%. The Alpha, Delta, Mu, and Omicron variants form one cluster, and GH/490R, Beta, and Gamma form another. Lambda appears to have emerged directly from Wuhan-Hu-1. In the first cluster, Omicron forms a long branch from the other members of the group.

4. Discussion

The number of whole-genome sequence data for SARS-CoV-2 submitted to international databases poses a major computational challenge to obtain an indication of the possible impact of each strain before clinical and experimental data are available, especially in resource-limited countries. At the time of writing of this paper, the number of submitted whole-genome SARS-CoV-2 sequences was approximately 6.5 million. Here, we offer a simple bioinformatic protocol for gene mining and predicting the possible biological meaning of genetic changes in strains. The GISAID initiative has enabled data mining by providing a variant tag for each submitted sequence. We proposed a guideline of the probable biological function of each residue

TABLE 1: Polymorphic amino acids residues of spike protein of SARS-CoV-2 Wuhan Hu-1 and all variants with possible biological function.

Amino acid position*	Wuhan Hu-1	SARS-CoV-2 variant							GH/490R	Known Function/probable biological impact**
		Alpha	Beta	Gamma	Delta	Lambda	Mu	Omicron		
9	P								L	SP
18	L		F	F						NTD
19	T				R					NTD; GML
20	T			N						NTD; AGM
26	P			S						NTD; IdA; PCE
67	A							V		NTD
69	H	Del						Del		NTD
70	V	Del						Del		NTD; PCE
75	G					V				NTD
76	T					I				NTD; GML
80	D		A							NTD
95	T				I		I	I		NTD
136	C								Del	NTD; CRL
137	N								Del	NTD
138	D			Y					Del	NTD
139	P								Del	NTD
140	F								Del	NTD
141	L								Del	NTD
142	G				D			D	Del	NTD
143	V	Del						Del	Del	NTD
144	Del	Del	Del	Del	Del	Del	T	Del	Del	NTD
145	Y	V					S	Del	Del	NTD
146	Y						N	Del		NTD
154	M				T					NTD
157	E				G					NTD
158	F				Del					NTD
159	R				Del					NTD
191	R			S						NTD; AGM
212	N							I		NTD; PCE
213	L							V		NTD; PCE
214	V							R		NTD; PCE
215	R							E		NTD
216	Del	Del	Del	Del	Del	Del	Del	P	Del	NTD
217	Del	Del	Del	Del	Del	Del	Del	E	Del	NTD
218	D		G						G	NTD
244	L		Del							NTD; IdB
245	L		Del							NTD; IdB
246	A		Del							NTD; IdB
248	H								P	NTD; IdB
249	R					N				NTD; AGM
250	S					Del				NTD; PCE
251	Y					Del				NTD; PCE
252	L					Del				NTD; PCE
253	T					Del				NTD; PCE
254	P					Del				NTD; PCE
255	G					Del				NTD; PCE
256	D					Del				NTD; PCE
342	G							D		RBD; IdD
349	R						K			RBD; IdD
374	S							L		RBD
376	S							P		RBD
378	S							F		RBD; IdE/He4
420	K		N	T				N		RBD; He5
443	N							K		RBD; RBS
449	G							S		RBD; RBS
455	L				R	Q				RBD; RBS; IdF
480	S							N		RBD; RBS
481	T				K			K		RBD; RBS

TABLE 1: Continued.

Amino acid position*	Wuhan Hu-1	SARS-CoV-2 variant							GH/490R	Known Function/probable biological impact**
		Alpha	Beta	Gamma	Delta	Lambda	Mu	Omicron		
487	E		K	K			K	A	K	RBD; RBS; IdG
493	F					S				RBD; RBS; IdG
496	Q							R		RBD; RBS; IdG
499	G							S		RBD; RBS; IdG
501	Q							R		RBD; RBS; IdG
504	N	Y	Y	Y			Y	Y		RBD; RBS
508	Y							H		RBD
550	T							K		IdH/He6-7
573	A	D								IdH/He6-7
617	D	G	G	G	G	G	G	G	G	IdH/He6-7
658	H			Y				Y		
678	Q					H				S1/S2-CS
682	N							K		S1/S2-CS; PCE
684	P	H			R		H	H	R	S1/S2-CS; PCE
693	Q		H							S1/S2-CS; PCE
704	A		V							
719	T	I								GML
767	N							K		He9-11
799	D							Y		FP
858	N							K		
862	T					N				
953	D				N		N			HR1
957	Q							H		HR1
972	N							K		HR1
984	L							F		HR1; PCE
985	S	A								HR1; PCE
1023	A						S			
1030	T			I						
1121	D	H								
1179	V			F						HR2; TM

*The positions were determined after alignment of all variants as available at supplementary material. Numbering 1–143 is equal to residues no. 1–143 of Wuhan-Hu-1. Number 144–215 is Wuhan-Hu-1 plus 1. Number >215 become Wuhan-Hu-1 plus 3; **SP: signal peptide; NTD: N-terminal domain of S1; S1/S2 CS: S1/S2 cleavage site; RBD: receptor binding domain; RBS: receptor binding site; FP: fusion peptide; HR1 or HR2: heptad repeat 1 or 2; TM: transmembrane; IdA, IdB, IdC, IdD, IdE/He4, IdF, IdG, IdH/He6-7, IdI/He12-13, He1, He2-3, He5, He8, He9-11, He14, He15, and He16: corresponding linear epitopes as described in Supplementary Material 1; PCE: probable conformational epitopes; GML: glycosylation motive loss; AGM: additional glycosylation motive; CRL: cysteine residue loss.

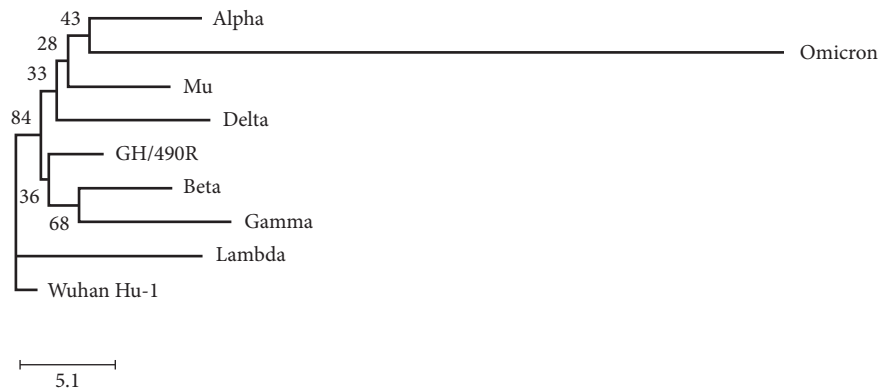


FIGURE 1: Evolutionary relationships of variants of SARS-CoV-2. The evolutionary history was inferred using the neighbor-joining method [13]. The evolutionary distances were computed using the Kimura 2-parameter method [14]. Evolutionary analyses were conducted in MEGA X [12]. The tree was rooted to Wuhan-Hu-1.

in the spike protein of SARS-CoV-2 based on current knowledge which could be adjusted following the fast flow of upcoming research reports.

It is generally believed that the phenotypic nature of viruses is mostly polygenic, whereby the entire genetic composition of strains determines the biology of the virus. This should also be true for coronaviruses, including SARS-CoV-2. The mechanism of pathogenesis of the Middle East severe acute respiratory syndrome coronavirus (MERS-CoV) mainly occurs through interaction between spike and cellular receptors, papain-like protease PLpro, and accessory proteins such as p4a and membrane *M* protein [15]. For SARS-CoV-2, various nonstructural proteins, such as PLpro [16] and various accessory proteins [5], have been described as contributing to virus biology and pathogenesis. However, focusing on the spike gene is also important, as a body of literature provides evidence on the key functions of this protein. The S1 domain, which carries major antigenic determinants, mediates receptor recognition and viral attachment to initiate host cell entry [17]. The NTD domain contributes to the host range [18]. Binding of the receptor-binding domain to receptors initiates infection [19], and the S2 domain mediates membrane fusion [17, 20, 21].

Various polymorphic amino acids are located mostly in the S1 domain, downstream from the S1/S2 cleavage site. The number of polymorphic amino acids at this site is 73, with only 15 in the S2 domain. We believe that as the S1 domain of the spike is located on the surface of the virion, thus allowing many substitutions, but that the S2 domain must be conserved to preserve virus integrity. Glycosylation and cysteine residues are also crucial for maintaining virus integrity. Our data show that glycosylation motif loss occurred once in the Delta and Lambda variants and that additional glycosylation motif gain occurred twice in Gamma and once in Lambda. Cysteine residue loss (CRL) occurred only once in the GH/490R variant.

The newest variant, Omicron, which was recently identified, exhibits nine amino acid changes in the RBD, whereas other variants are more conserved, with only one or two residue differences from Wuhan-Hu-1. Therefore, it is plausible to expect biological changes in Omicron that differ from those of other variants as well as the original strain. Tracking variant occurrence on the GISAID website revealed a sharp global increase from under 1% on November 29, 2021, to 50% of circulating strains on December 29, 2021. Higher transmissibility is evident.

The Omicron linear and conformational epitopes diverge most from the Covid-19 origin strain Wuhan-Hu-1, with ten and seven amino acid substitutions, respectively. Seven amino acid substitutions at conformational epitopes also occur in the Lambda variant. The other variants have only 2–6 amino acid changes in linear and 1–3 substitutions in conformational epitopes of spike. The identified SARS-CoV-2 spike epitopes consist of at least 10–20 residues or longer. As MHC-1-presenting B- and T cell epitopes are limited to 9–11 residues and MHC-II-presenting epitopes are limited to 9–22 [22], the identified epitopes of the spike protein of SARS-CoV-2 need to be refined. Moreover, due to the multiple epitopes of more than 20, a slight change at

some epitopes while others are conserved should not lead to a significant reduction in the effectiveness of existing vaccines. Reports on the reduced efficacy of vaccines against variants based on in vitro experiments [23] should not cause concern, as they might not significantly reduce vaccine efficacy in vivo. Indeed, the immune system consists of an array of components, such as cytokines, complement activation, and macrophage opsonization [24–26], as already reported in SARS-CoV-2 [27–30]. Moreover, cellular mediated immunity must pose critical role in a complete immune protection in SARS-CoV-2 [31], which is not involved in an in vitro neutralization testing.

The S1/S2 cleavage site changes from nonbasic Q/N/P to basic H/R/K amino acids might be critical to the biology of the virus, especially Omicron. Although other variants show a single amino acid change, Omicron exhibits two. Moreover, all variants carry more basic S1/S2 cleavage sites, and Omicron has the most basic S1/S2 cleavage site. It is well-documented for influenza viruses that the polybasic cleavage site allows for ubiquitous cellular protease activation for the virus to initiate infection [32–34]. The cleavage site of the origin of SARS-CoV-2 is indeed polybasic with an NH-PRRAR-COOH motif. A P681H change occurred in Alpha, Mu, and Omicron, with P681R in delta and GH/490R. A more basic cleavage site was acquired in Omicron due to the N679K substitution. For the Delta variant, the more basic cleavage site might have contributed to its transmissibility and clinical outcomes [35, 36]. Therefore, it is plausible to expect higher transmissibility of Omicron as its cleavage site is more basic than that of the delta variant. However, its clinical consequences among nonimmune people, who are unvaccinated or have not experienced natural infection, are expected to be reported soon.

Despite the protein changes of Omicron, it has a relatively conserved S2 domain. Omicron shows amino acid alterations in the fusion peptide and HR1, with one and three substitutions, respectively, which might have an impact on fusion capability. As described previously, these domains mediate membrane fusion in infected cells [17, 20, 21].

Phylogenetic analysis showed that the Omicron forms a long branch from the other members of the cluster. This phenomenon is most frequently observed due to isolation and having no known close relatives [37]. The ancestor of this variant might have been circulating without notice, or the number of genome sequences from the area of circulation was limited. Another explanation is that dramatic changes might have occurred shortly before its identification.

We expect some biological changes due to amino acid substitution of the spike protein of SARS-CoV-2, especially for omicron. The most significant number of amino acid substitutions occurred in Omicron, with divergence in the RBS, epitopes, S1/S2 cleavage site, fusion peptide, and HR1. The sharp global increase in dominance of 50% of circulating viruses shows evidence of their high community transmissibility. As multiple epitopes of more than 20 residues exist on spike, a slight change in some epitopes while others are conserved should not lead to a significant reduction in

existing vaccine effectiveness. The emergence of Omicron with dramatic changes and an unknown recent ancestor should draw attention from the global community for resource mobilization.

In conclusion, the proposed guideline could give an immediate insight into the probable biological nature of any variant of SARS-Cov-2. The Omicron diverged the farthest from the original pandemic strain Wuhan-Hu-1 with divergence in the RBS, epitopes, S1/S2 cleavage site, fusion peptide, and HR1. Therefore, we expect different epidemiological and clinical pattern of Omicron cases. On the vaccine efficacy, slight changes in some epitopes while others are conserved should not lead to a significant reduction in existing vaccine effectiveness.

Abbreviations

AGM:	Additional glycosylation motive
CRL:	Cysteine residue loss
FP:	Fusion peptide
GML:	Glycosylation motive loss
HR1 or HR2:	Heptad repeat 1 or 2
NTD:	N-terminal domain of S1
PCE:	Probable conformational epitopes
RBD:	Receptor binding domain
RBS:	Receptor binding site
S1/S2 CS:	S1/S2 cleavage site
SARS-CoV-2:	Severe acute respiratory syndrome coronavirus 2
SP:	Signal peptide
TM:	Transmembrane
VOC:	Variant of concern
VOI:	Variant of interest
VUM:	Variant under monitoring.

Data Availability

All the data are provided in the text as well as in the supplementary materials.

Additional Points

Statement of significance of the study. The number of whole-genome sequence data for SARS-CoV-2 submitted to international databases poses a major computational challenge to obtain an indication of the possible impact of each strain before clinical and experimental data are available, especially in resource-limited countries. Here, we offer a simple bioinformatic protocol for gene mining and predicting the possible biological meaning of genetic changes in strains. We proposed a guideline of the probable biological function of each residue in the spike protein of SARS-Cov-2 based on current knowledge which could be adjusted following the fast flow of upcoming research reports. Applying the guideline to representatives of various SARS-CoV-2 variants, it is remarkable that polymorphic amino acids are located mostly on the amino-terminal side of the S1/S2

cleavage site. The Omicron variant diverges from the others at including the receptor-binding site (RBS), epitopes, S1/S2 cleavage site, fusion peptide, and heptad repeat 1. The unique RBS and S1/S2 cleavage sites could contribute to the sharp global increase of Omicron cases. However, a slight change in some epitopes while others are conserved should not lead to a significant reduction in the effectiveness of existing vaccines.

Conflicts of Interest

The authors have declared no conflicts of interest.

Acknowledgments

This study was supported by Research, Technology and Higher Education (RISTEKDIKTI) of Indonesia through World Class Research. The English expression of the manuscript has been copy-edited by a global professional copy-editing service.

Supplementary Materials

Supplementary Materials. The probable biological function of each residue was annotated using the guidelines shown in Supplementary Material 1. The dataset containing the representative of each variant used is available in Supplementary Material 2. (*Supplementary Materials*)

References

- [1] F. Wu, S. Zhao, B. Yu et al., "A new coronavirus associated with human respiratory disease in China," *Nature*, vol. 579, no. 7798, pp. 265–269, 2020.
- [2] Y. Jin, H. Yang, W. Ji et al., "Virology, epidemiology, pathogenesis, and control of COVID-19," *Viruses*, vol. 12, no. 4, 2020.
- [3] P. R. Bhatt, A. Scaiola, G. Loughran et al., "Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome," *Science*, vol. 372, no. 6548, pp. 1306–1313, 2021.
- [4] F. K. Yoshimoto, "The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19," *The Protein Journal*, vol. 39, no. 3, pp. 198–216, 2020.
- [5] N. Redondo, S. Zaldívar-López, J. J. Garrido, and M. Montoya, "SARS-CoV-2 accessory proteins in viral pathogenesis: knowns and unknowns," *Frontiers in Immunology*, vol. 12, Article ID 708264, 2021.
- [6] P. S. Masters, "The molecular biology of coronaviruses," *Advances in Virus Research*, vol. 66, pp. 193–292, 2006.
- [7] W. Tian, D. Li, N. Zhang et al., "O-glycosylation pattern of the SARS-CoV-2 spike protein reveals an "O-Follow-N" rule," *Cell Research*, vol. 31, no. 10, pp. 1123–1125, 2021.
- [8] J. K. Millet and G. R. Whittaker, "Host cell proteases: critical determinants of coronavirus tropism and pathogenesis," *Virus Research*, vol. 202, pp. 120–134, 2015.
- [9] B.-z. Zhang, Y.-f. Hu, L.-l. Chen et al., "Mining of epitopes on spike protein of SARS-CoV-2 from COVID-19 patients," *Cell Research*, vol. 30, no. 8, pp. 702–704, 2020.

- [10] J. He, F. Huang, J. Zhang et al., "Vaccine design based on 16 epitopes of SARS-CoV-2 spike protein," *Journal of Medical Virology*, vol. 93, no. 4, pp. 2115–2131, 2021.
- [11] W. T. Harvey, A. M. Carabelli, B. Jackson et al., "SARS-CoV-2 variants, spike mutations and immune escape," *Nature Reviews Microbiology*, vol. 19, no. 7, pp. 409–424, 2021.
- [12] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, and X. Mega, "Mega X: molecular evolutionary genetics analysis across computing platforms," *Molecular Biology and Evolution*, vol. 35, no. 6, pp. 1547–1549, 2018.
- [13] N. Saitou and M. Nei, "The neighbor-joining method: a new method for reconstructing phylogenetic trees," *Molecular Biology and Evolution*, vol. 4, no. 4, pp. 406–425, 1987.
- [14] M. Kimura, "A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences," *Journal of Molecular Evolution*, vol. 16, no. 2, pp. 111–120, 1980.
- [15] S. Skariyachan, S. B. Challapilli, S. Packirisamy, S. T. Kumargowda, and V. S. Sridhar, "Recent aspects on the pathogenesis mechanism, animal models and novel therapeutic interventions for middle East respiratory syndrome coronavirus infections," *Frontiers in Microbiology*, vol. 10, p. 569, 2019.
- [16] D. Shin, R. Mukherjee, D. Grewe et al., "Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity," *Nature*, vol. 587, no. 7835, pp. 657–662, 2020.
- [17] A. C. Walls, Y.-J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, and D. Veasler, "Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein," *Cell*, vol. 183, no. 6, p. 1735, 2020.
- [18] M. Sironi, S. E. Hasnain, B. Rosenthal et al., "SARS-CoV-2 and COVID-19: a genetic, epidemiological, and evolutionary perspective," *Infection, Genetics and Evolution*, vol. 84, Article ID 104384, 2020.
- [19] W. Tai, L. He, X. Zhang et al., "Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine," *Cellular & Molecular Immunology*, vol. 17, no. 6, pp. 613–620, 2020.
- [20] S. Xia, Y. Zhu, M. Liu et al., "Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein," *Cellular & Molecular Immunology*, vol. 17, no. 7, pp. 765–767, 2020.
- [21] C. M. Petit, J. M. Melancon, V. N. Chouljenko et al., "Genetic analysis of the SARS-coronavirus spike glycoprotein functional domains involved in cell-surface expression and cell-to-cell fusion," *Virology*, vol. 341, no. 2, pp. 215–230, 2005.
- [22] J. L. Sanchez-Trincado, M. Gomez-Perosanz, and P. A. Reche, "Fundamentals and methods for T- and B-cell epitope prediction," *Journal of Immunology Research*, vol. 2017, Article ID 2680160, 14 pages, 2017.
- [23] S. A. Madhi, V. Baillie, C. L. Cutland et al., "Efficacy of the ChAdOx1 nCoV-19 covid-19 vaccine against the B.1.351 variant," *New England Journal of Medicine*, vol. 384, no. 20, pp. 1885–1898, 2021.
- [24] G. Arango Duque and A. Descoteaux, "Macrophage cytokines: involvement in immunity and infectious diseases," *Frontiers in Immunology*, vol. 5, p. 491, 2014.
- [25] F. R. Toapanta and T. M. Ross, "Complement-mediated activation of the adaptive immune responses: role of C3d in linking the innate and adaptive immunity," *Immunologic Research*, vol. 36, no. 1-3, pp. 197–210, 2006.
- [26] J. R. Dunkelberger and W.-C. Song, "Complement and its role in innate and adaptive immune responses," *Cell Research*, vol. 20, no. 1, pp. 34–50, 2010.
- [27] V. Shinde, S. Bhikha, Z. Hoosain et al., "Efficacy of NVX-CoV2373 covid-19 vaccine against the B.1.351 variant," *New England Journal of Medicine*, vol. 384, no. 20, pp. 1899–1909, 2021.
- [28] O. T. Ranzani, M. D. T. Hitchings, M. Dorion et al., "Effectiveness of the CoronaVac vaccine in older adults during a gamma variant associated epidemic of covid-19 in Brazil: test negative case-control study," *BMJ*, vol. 374, Article ID n2015, 2021.
- [29] G. N. Sapkal, P. D. Yadav, R. Ella et al., "Inactivated COVID-19 vaccine BBV152/COVAXIN effectively neutralizes recently emerged B.1.1.7 variant of SARS-CoV-2," *Journal of Travel Medicine*, vol. 28, no. 4, 2021.
- [30] P. Tang, M. R. Hasan, H. Chemaitelly et al., "BNT162b2 and mRNA-1273 COVID-19 vaccine effectiveness against the SARS-CoV-2 delta variant in qatar," *Nature Medicine*, vol. 27, no. 12, pp. 2136–2143, 2021.
- [31] J. T. Sandberg, R. Varnaitė, W. Christ et al., "SARS-CoV-2-specific humoral and cellular immunity persists through 9 months irrespective of COVID-19 severity at hospitalisation," *Clinical & translational immunology*, vol. 10, no. 7, Article ID e1306, 2021.
- [32] N. Nao, J. Yamagishi, H. Miyamoto et al., "Genetic predisposition to acquire a polybasic cleavage site for highly pathogenic avian influenza virus hemagglutinin," *mBio*, vol. 8, no. 1, 2017.
- [33] D. Scheibner, R. Ulrich, O. I. Fatola et al., "Variable impact of the hemagglutinin polybasic cleavage site on virulence and pathogenesis of avian influenza H7N7 virus in chickens, turkeys and ducks," *Scientific Reports*, vol. 9, no. 1, Article ID 11556, 2019.
- [34] R. Webby, D. Perez, J. Coleman et al., "Responsiveness to a pandemic alert: use of reverse genetics for rapid development of influenza vaccines," *The Lancet*, vol. 363, pp. 1099–1103, 2004.
- [35] Y. Liu and J. Rocklöv, "The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus," *Journal of Travel Medicine*, vol. 28, no. 7, 2021.
- [36] B. H. Ryu, S. I. Hong, S. J. Lim et al., "Clinical features of adult COVID-19 patients without risk factors before and after the nationwide SARS-CoV-2 B.1.617.2 (Delta)-variant outbreak in Korea: experience from Gyeongangnam-do," *Journal of Korean Medical Science*, vol. 36, no. 49, p. e341, 2021.
- [37] A. E. Gorbalenya and C. Lauber, "Phylogeny of viruses," *Reference Module in Biomedical Sciences*, pp. 1–7, 2017.